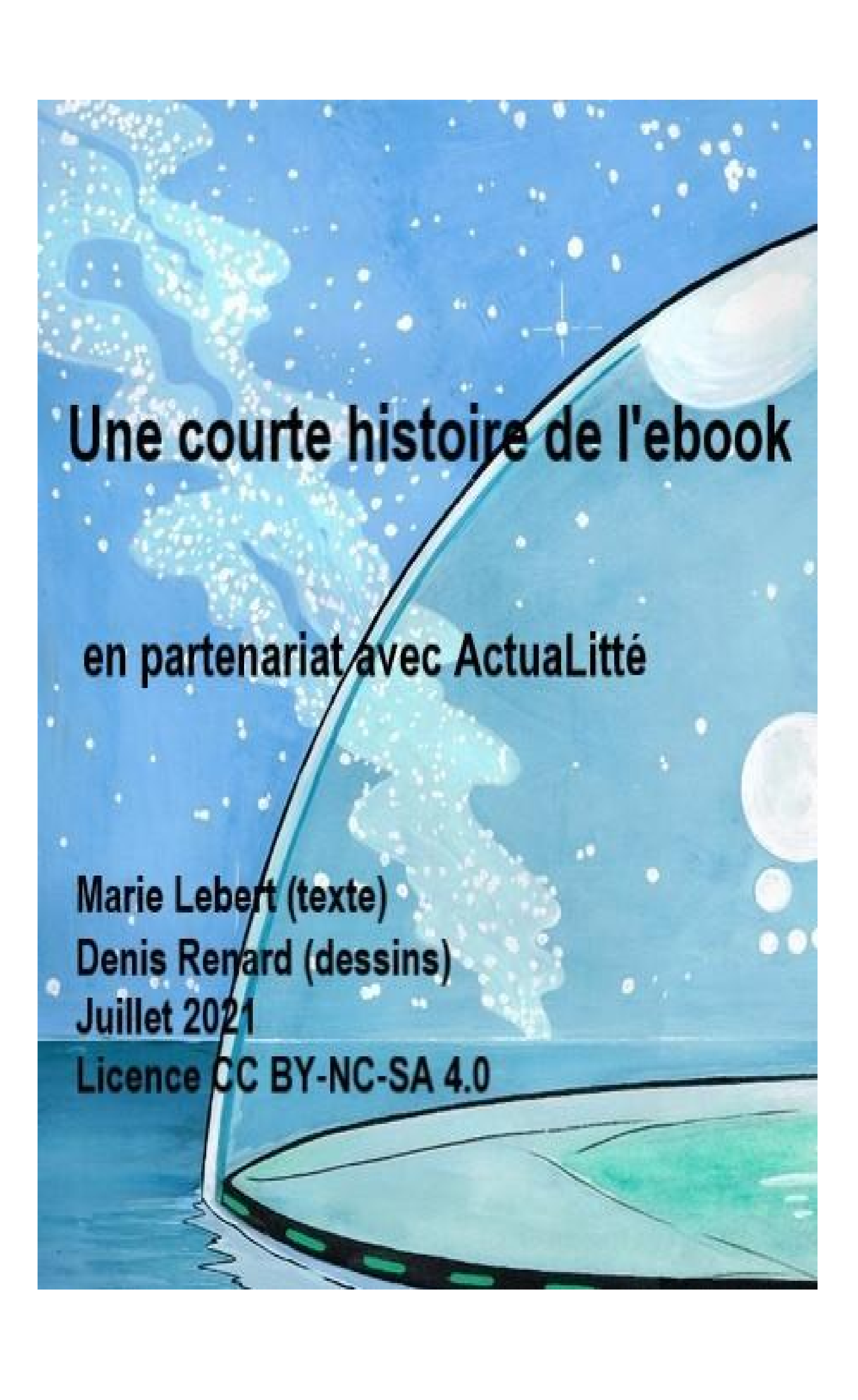
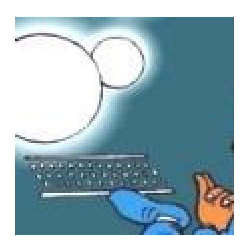

Une courte histoire de l'ebook
en partenariat avec ActuaLitté
Marie Lebert (texte)
Denis Renard (dessins)
Juillet 2021
Licence CC BY-NC-SA 4.0



Table des matières

@@@@@@@

 Remerciements

Merci à Nicolas Gary, directeur de publication d'ActuaLitté, pour la diffusion de ce livre sous forme d'articles.

Merci à Denis Renard pour ses beaux dessins.

Merci à tous les professionnels du livre interviewés de par le monde pendant près de vingt ans pour mes articles et livres sur le sujet.

Merci au Projet Gutenberg de nous offrir plus de 65.000 livres numériques gratuits de qualité dans 60 langues pour ses 50 ans.

Merci aux volontaires de Distributed Proofreaders pour la révision de ces livres à deux reprises.

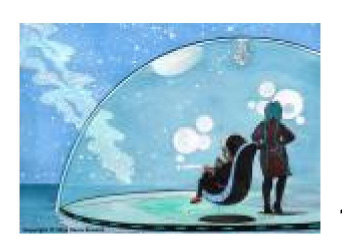
Merci à l'Internet Archive pour ses Community Texts ouverts à tous.

Merci à tous les bibliothécaires de gérer au mieux tant de documents.

Merci à tous les traducteurs et traductrices professionnels, trop souvent oubliés.

Marie Lebert
marie.lebert@gmail.com

@@@@@@@



1. Le Projet Gutenberg, un projet visionnaire

Le Projet Gutenberg est fondé par Michael Hart en juillet 1971 pour créer des versions numériques d'œuvres littéraires du domaine public et les diffuser gratuitement dans le monde entier. D'abord considéré comme complètement irréaliste, ce projet prend son envol avec l'invention du web en 1991. Les ebooks sont tous relus à deux reprises sur la plateforme Distributed Proofreaders avec le scan de la version imprimée en vis-à-vis. Le Projet Gutenberg fête ses 50 ans en juillet 2021 avec plus de 65.000 ebooks en 60 langues et des dizaines de milliers de téléchargements par jour.

***

Les débuts du Projet Gutenberg

Alors étudiant à l'Université de l'Illinois (États-Unis), Michael Hart se voit attribuer quelques centaines de milliers d'heures de «temps machine» (ce qui correspond à quelques millions de dollars) dans le laboratoire informatique (Materials Research Lab) de son université.

Le 4 juillet 1971, jour de la fête nationale, il saisit «The United States Declaration of Independence» sur le clavier de son terminal. En caractères majuscules, puisque les caractères minuscules n'existent pas encore. Le texte électronique représente 5 Ko (kilo-octets). Puis il diffuse un message aux 100 usagers du réseau pré-internet pour indiquer où le fichier est stocké, suite à quoi ce fichier est téléchargé par six personnes. Le Projet Gutenberg est né.

Dans la foulée, Michael Hart décide de se consacrer à la recherche d'œuvres littéraires disponibles en bibliothèque, à la numérisation de celles-ci (saisie sur un clavier page après page) et au stockage des textes électroniques. Peu après, il définit la mission du Projet Gutenberg, à savoir mettre gratuitement à la disposition de tous, par voie électronique, le plus grand nombre possible d'oeuvres littéraires.

D'abord considéré comme complètement irréaliste, ce projet prend son envol en 1991 avec l'apparition du web, ce qui va faciliter la circulation des textes électroniques et le recrutement des volontaires.

Michael Hart explique en août 1998 dans un entretien par courriel: «Nous considérons le texte électronique comme un nouveau médium, sans véritable relation avec le papier. Le seul point commun est que nous diffusons les mêmes œuvres, mais je ne vois pas comment le papier peut concurrencer le texte électronique une fois que les gens y sont habitués, particulièrement dans les établissements d'enseignement.»

Au lieu d’être un ensemble de pages brochées ou reliées, le livre devient un texte électronique en continu au format ASCII, avec des lettres capitales pour les termes en italique, en gras et soulignés de la version imprimée. Le format ASCII (American Standard Code for Information Interchange) est le format le plus simple et le plus répandu et il peut être lu sur tout ordinateur, toute plateforme et tout logiciel. Il sera plus tard remplacé par l’Unicode, qui peut prendre en compte toutes les langues de la planète.

La plateforme Distributed Proofreaders

Le Projet Gutenberg trouve un nouveau souffle avec la création de la plateforme Distributed Proofreaders pour partager la relecture des livres numérisés entre des milliers de volontaires. Conçu en octobre 2000 par Charles Franks pour aider à la numérisation des livres du domaine public, Distributed Proofreaders devient rapidement la principale source du Projet Gutenberg.

L’idée est de permettre la relecture partagée de livres scannés à partir d'une version imprimée puis convertis au format texte par un logiciel OCR (Optical Character Recognition), fiable à 99% dans le meilleur des cas, ce qui représente donc quelques erreurs par page. Les fichiers obtenus sont fragmentés en pages pouvant être relues par différents volontaires. Le scan de la page imprimée et le fichier OCR apparaissent en vis-à-vis à l’écran. Les volontaires n'ont aucun quota à respecter. À titre indicatif, il est suggéré de relire une page par jour. Les livres sont tous relus deux fois de bout en bout par des volontaires différents.

Distributed Proofreaders est officiellement affilié au Projet Gutenberg en 2002, puis devient une entité séparée avec la Distributed Proofreaders Foundation en mai 2006. Distributed Proofreaders compte 10.000 titres numérisés et relus en décembre 2006, 20.000 titres en avril 2011 et plus de 40.000 titres en juin 2021.

La philosophie du Projet Gutenberg

La structure administrative et financière du Projet Gutenberg se limite au strict minimum, avec une devise qui tient en trois mots: «Less is more» (Le plus est le moins).

Le but est d’assurer la pérennité du projet indépendamment des crédits, des coupures de crédits et des priorités culturelles, financières et politiques du moment. Pas de pression possible donc par le pouvoir et par l’argent. Et respect à l’égard des volontaires, qui sont assurés de voir leur travail utilisé pendant des décennies, si ce n’est pour plusieurs générations. Le suivi régulier du projet est assuré grâce à une lettre d’information hebdomadaire et mensuelle, des forums de discussion, des wikis et des blogs.

En juillet 2011, quarante ans après les débuts du Projet Gutenberg et peu avant son décès, Michael Hart se définit toujours comme un fou de travail dédiant toute sa vie à son projet, qu’il voit comme étant à l’origine d’une révolution néo-industrielle. Il se définit aussi comme altruiste, pragmatique et visionnaire. Après avoir été traité de toqué pendant des années, il force maintenant le respect.

Michael Hart précise souvent dans ses écrits que, si Gutenberg a permis à chacun d'avoir ses propres livres (jusque-là réservés à une élite) pour un coût relativement modique, le Projet Gutenberg permet à chacun d'avoir une bibliothèque gratuite (jusque-là réservée à une collectivité) sur un support qu'on peut glisser dans sa poche ou porter en pendentif autour du cou. Les collections du Projet Gutenberg ont la taille d'une bibliothèque publique de quartier, mais cette fois disponible sur le web, téléchargeable par tous et indéfiniment reproductible.

Au fil des ans, sous la houlette de Greg Newby, son CEO, la mission du Projet Gutenberg reste la même, à savoir produire et diffuser des livres numériques gratuits de qualité, et favoriser ainsi la lecture et la culture pour tous dans le monde entier.

@@@@@@@



2. Les grandes dates du Projet Gutenberg

«Nous considérons le texte électronique comme un nouveau médium, sans véritable relation avec le papier. Le seul point commun est que nous diffusons les mêmes œuvres, mais je ne vois pas comment le papier peut concurrencer le texte électronique une fois que les gens y sont habitués, particulièrement dans les établissements d'enseignement.» (Michael Hart, fondateur du Projet Gutenberg et père de l'ebook)

***

[année-mois]

1971-07 > Michael Hart saisit "The United States Declaration of Independence" (eBook #1) le 4 juillet 1971 et mentionne le fichier auprès des 100 usagers du réseau pré-internet. Six personnes le téléchargent. Le Projet Gutenberg est né.

1972 > L'eBook #2 est "The United States Bill of Rights".

1973 > L'eBook #5 est "The United States Constitution".

1974 > L'internet fait ses débuts et va bientôt changer le monde… y compris le monde du livre.

1974 à 1988 > Michael Hart saisit patiemment la Bible et plusieurs pièces de Shakespeare, sous forme de petits fichiers, d'abord seul puis avec quelques volontaires.

1989-08 > L'eBook #10 est "The King James Bible".

1991-01 > L'eBook #11 est "Alice's Adventures in Wonderland" ("Alice au pays des merveilles") de Lewis Carroll, qui tient sur une disquette de l'époque.

1991-06 > L'eBook #16 est "Peter Pan" de J. M. Barrie, qui tient lui aussi sur une disquette.

1991 > Les débuts du web vont faciliter la diffusion des ebooks et le recrutement des volontaires.

1991 > Numérisation (saisie d'un livre page après page sur un clavier d'ordinateur) d'un livre par mois.

1992 > Numérisation de deux livres par mois.

1993 > Numérisation de quatre livres par mois.

1993-12 > Création de trois grands secteurs: Light Literature, Heavy Literature, Reference Literature.

1994-01 > L'eBook #100 est "The Complete Works of William Shakespeare".

1994 > Numérisation de huit livres par mois.

1995 > Numérisation de 16 livres par mois.

1996 et 1997 > Numérisation de 32 livres par mois.

1997-08 > L'eBook #1000 est "La Divina Commedia" de Dante, en italien.

1998 à 2000 > Numérisation de 36 livres par mois.

1999-05 > L'eBook #2000 est "Don Quijote" de Cervantes, en espagnol.

2000 > Création de la Project Gutenberg Literary Archive Foundation (PGLAF).

2000-10 > La plateforme Distributed Proofreaders est créée par Charles Franks pour permettre la relecture partagée des ebooks suite au scan OCR des livres imprimés.

2000-12 > L'eBook #3000 est le troisième volume de "À l'ombre des jeunes filles en fleurs" de Marcel Proust, en français.

2001-08 > Création du Project Gutenberg of Australia.

2001-10 > L'eBook #4000 est "The French Immortals Series", dans sa traduction anglaise.

2001: Numérisation (scan OCR puis relecture) de 104 livres par mois.

2002-04 > L'eBook #5000 est "The Notebooks of Leonardo da Vinci", en anglais.

2002 > Numérisation de 203 livres par mois.

2003-08 > Édition d'un CD Best of Gutenberg contenant 600 ebooks.

2003-09 > Lancement de la section Project Gutenberg Audio eBooks.

2003-10 > Les collections ont doublé en 18 mois, passant de 5.000 à 10.000 ebooks.

2003-10 > L'eBook #10000 est "The Magna Carta".

2003-12 > Édition du premier DVD du Projet Gutenberg, qui contient 9.400 ebooks (soit pratiquement toute la collection) à destination des bibliothèques et des écoles.

2003 > Numérisation de 348 livres par mois.

2004-01 > Lancement du Projet Gutenberg Europe par le Projet Rastko.

2004-02 > Voyage de Michael Hart en Europe avec une conférence au siège de l'UNESCO (Paris), une visite au Parlement européen (Bruxelles) et une rencontre avec le Projet Rastko (Belgrade).

2004-10 > 5.000 titres ont été scannés et relus par Distributed Proofreaders.

2004 > Numérisation de 338 livres par mois.

2005-01 > L'eBook #15000 est "The Life of Reason" de George Santayana.

2005-06 > Le Projet Gutenberg compte 16.000 ebooks.

2005-10 > Distributed Proofreaders fête ses cinq ans.

2006-01 > Lancement de la section Project Gutenberg PrePrints.

2006-05 > Création de la Distributed Proofreaders Foundation.

2006-07 > Édition d'un nouveau DVD incluant 17.000 ebooks à destination des bibliothèques et des écoles.

2006-12 > L'eBook #20000 est "Twenty Thousand Leagues Under the Sea" ("Vingt mille lieues sous la mer") de Jules Verne, un audiobook en anglais.

2006 > Numérisation de 345 livres par mois.

2007-03 > 10.000 titres ont été scannés et relus par Distributed Proofreaders.

2007-07 > Création du Project Gutenberg Canada (PGC).

2008-04 > L'eBook #25000 est "English Book Collectors" de William Younger Fletcher.

2010-10 > Distributed Proofreaders fête ses 10 ans avec 18.000 titres scannés et relus pour le Projet Gutenberg.

2011-07 > Le Projet Gutenberg fête ses 40 ans et comprend maintenant 36.000 ebooks.

2011-09 > Michael Hart – fondateur du Projet Gutenberg et père de l'ebook – décède dans l'Illinois à l'âge de 64 ans après avoir consacré toute sa vie à son projet.

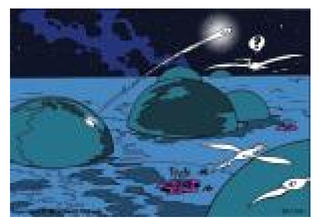
2015-07 > 30.000 titres ont été scannés et relus par Distributed Proofreaders.

2015-09 > Le Projet Gutenberg offre 50.000 ebooks, surtout des oeuvres littéraires du domaine public.

2020-10 > Distributed Proofreaders fête ses 20 ans avec près de 40.000 titres scannés et relus pour le Projet Gutenberg.

2021-07 > Le Projet Gutenberg fête ses 50 ans le 4 juillet 2021 avec plus de 65.000 ebooks dans 60 langues et dans plusieurs formats (dont l'EPUB et le Kindle), et plusieurs dizaines de milliers de téléchargements par jour.

@@@@@@@



3. Le PDF, format pionnier lancé par Adobe

Peu après les débuts du web en 1990, la société Adobe lance en juin 1993 le format PDF (Portable Document Format), tout comme l'Acrobat Reader (gratuit, pour lire les PDF) et l'Adobe Acrobat (payant, pour créer les PDF).

***

Un standard de diffusion

Le but du format PDF est de figer les documents numériques dans une présentation donnée, pour conserver la présentation originale du document source, quelle que soit la plateforme utilisée pour le créer et pour le lire.

Au fil des ans, le format PDF devient un standard de diffusion des documents électroniques. L'Acrobat Reader est progressivement disponible dans plusieurs langues et pour diverses plateformes (Windows, Mac, Linux).

En 2000, Adobe passe un partenariat avec les grandes librairies en ligne Barnes & Noble.com (an août 2000) et Amazon.com (en novembre 2000) pour que celles-ci proposent des titres lisibles sur l'Acrobat Reader dans leur eBookStore.

Deux nouveaux logiciels

En janvier 2001, Adobe lance deux nouveaux logiciels.

Le premier logiciel, gratuit, est l'Acrobat eBook Reader, qui permet de lire les fichiers PDF de livres numériques sous droits (avec gestion des droits par l'Adobe Content Server). Ce logiciel permet d'ajouter des notes et des signets, de choisir l'orientation de lecture des livres (paysage ou portrait), ou encore de visualiser leur couverture dans une bibliothèque personnelle. Il utilise la technique d'affichage CoolType et comporte un dictionnaire intégré.

Le deuxième logiciel, payant, est l'Adobe Content Server, destiné aux éditeurs et distributeurs. Il s'agit d'un logiciel serveur de contenu assurant le conditionnement, la protection, la distribution et la vente sécurisée de livres numériques au format PDF.

Ce système de gestion des droits numériques (ou système DRM: Digital Rights Management) permet de contrôler l'accès aux livres numériques sous droits, et donc de gérer les droits d'un livre selon les consignes données par le gestionnaire des droits (qui est

souvent l'éditeur), par exemple en autorisant ou non l'impression ou le prêt. L'Adobe Content Server sera remplacé par l'Adobe LiveCycle Policy Server en novembre 2004.

En avril 2001, Adobe passe un deuxième partenariat avec Amazon, qui met en vente dans son eBookStore 2.000 livres numériques lisibles sur l'Acrobat eBook Reader. Ces livres sont des titres de grands éditeurs, des guides de voyages et des livres pour enfants.

L'Acrobat Reader s'enrichit d'une version PDA (Personal Digital Assistant), disponible pour le Palm Pilot en mai 2001 puis pour le Pocket PC en décembre 2001. En effet le grand public commence à lire sur PDA, suscitant l'inquiétude de certains professionnels du livre (et de certains ophtalmologues) qui trouvent l'écran du PDA trop petit, alors que les adeptes de la lecture sur PDA tentent de les convaincre du contraire.

L'Adobe Reader

En dix ans, entre 1993 et 2003, l'Acrobat Reader aurait été téléchargé 500 millions de fois.

En 2003, ce logiciel est désormais disponible dans de nombreuses langues et pour toute plateforme (Windows, Mac, Linux, Palm OS, Pocket PC, Symbian OS, etc.). 10% des documents présents sur l'internet seraient au format PDF. Des millions de fichiers PDF sont présents sur le web pour lecture et téléchargement ou bien transitent par courriel. Le format PDF est également le format de livre numérique le plus répandu.

En mai 2003, l'Acrobat Reader (version 5) fusionne avec l'Acrobat eBook Reader (version 2) pour devenir l'Adobe Reader, qui débute à la version 6 et permet de lire aussi bien les fichiers PDF standard que les fichiers PDF sécurisés des livres numériques sous droits.

Fin 2003, Adobe ouvre sa librairie en ligne, le Digital Media Store, avec les titres au format PDF de grands éditeurs tels que HarperCollins Publishers, Random House et Simon & Schuster, ainsi que les versions électroniques de journaux comme le New York Times et de magazines comme Popular Science.

Adobe lance aussi Adobe eBooks Central, un service permettant de lire, publier, vendre et prêter des livres numériques, et l'Adobe eBook Library, qui se veut un prototype de bibliothèque de livres numériques.

Après avoir été un format propriétaire, le format PDF devient un standard ouvert en juillet 2008. Il est publié en tant que norme ISO (Organisation internationale de normalisation) sous l'appellation ISO 32000-1:2008.

@@@@@@@



4. Gabriel, portail des bibliothèques nationales européennes

Mis en ligne en janvier 1997, Gabriel est un portail trilingue (anglais, allemand, français) offrant un point d'accès unique aux services internet des bibliothèques nationales européennes. Sans les patients efforts de Gabriel (et de l'European Library) pendant dix ans, la bibliothèque numérique européenne Europeana n'aurait peut-être jamais vu le jour dix ans plus tard.

***

Pourquoi ce nom Gabriel?

Gabriel est l'acronyme de «Gateway and Bridge to Europe's National Libraries».

On lit à l'époque sur le site que le choix de ce nom «rappelle également les travaux de Gabriel Naudé, dont l'"Advis pour dresser une bibliothèque" (Paris, 1627) est le premier travail théorique en Europe sur les bibliothèques et constitue ainsi un point de départ sur les bibliothèques de recherche modernes. Le nom Gabriel est aussi employé dans de nombreuses langues européennes et vient de l'Ancien Testament, Gabriel étant l'un des archanges, ou messager céleste. Il est également présent dans le Nouveau Testament et dans le Coran.»

Plus prosaïquement, le site offre en 1998 des liens hypertextes vers les services internet des 38 bibliothèques nationales participantes (Allemagne, Autriche, Belgique, Bulgarie, Danemark, Espagne, Estonie, Finlande, France, Grèce, Hongrie, Irlande, Islande, Italie, Lettonie, Liechtenstein, Lituanie, Luxembourg, Macédoine, Malte, Norvège, Pays-Bas, Pologne, Portugal, République slovaque, République tchèque, Roumanie, Royaume-Uni, San Marino, Suède, Suisse, Turquie et Vatican).

Les services internet sont très divers d'une bibliothèque à l'autre, avec une liste complète disponible par bibliothèque. Ces services sont par exemple des catalogues en ligne appelés aussi OPAC (Online Public Access Catalogues), des bibliographies nationales, des catalogues collectifs nationaux, des index de périodiques, des serveurs web et des gophers (à savoir des systèmes d'information à base de menus textuels à plusieurs niveaux). Une rubrique spécifique fait part des projets communs à plusieurs pays. La recherche sur Gabriel est possible par pays et par type de services.

Comment Gabriel voit-il le jour?

L’idée d’un projet commun aux bibliothèques nationales européennes naît lors de la réunion annuelle de la CENL (Conference of European National Librarians) en 1994 à Oslo (Norvège). Le projet débute par un tableau d’affichage électronique commun qui est régulièrement actualisé avec les projets internet en cours.

En mars 1995, une nouvelle réunion rassemble les représentants des bibliothèques nationales du Pays-Bas (Koninklijke Bibliotheek), du Royaume-Uni (British Library) et de Finlande (Helsinki University Library), qui élaborent un projet pilote.

Ils sont rejoints ensuite par leurs collègues des bibliothèques nationales d’Allemagne (Die Deusche Bibliothek), de France (Bibliothèque nationale de France) et de Pologne (Biblioteka Narodowa), le but étant de décrire leurs services et collections en tentant d’inciter d’autres bibliothèques nationales à participer au projet.

Lancé en septembre 1995, le premier site Gabriel est géré par la British Library, qui assure sa maintenance éditoriale, avec deux sites miroirs sur les serveurs des bibliothèques nationales des Pays-Bas et de Finlande.

La seconde étape se déroule sur onze mois, entre octobre 1995 et septembre 1996. Les bibliothèques nationales n’ayant pas participé à la phase pilote sont invitées à se joindre au projet et le nombre de bibliothèques utilisant Gabriel s’accroît.

Pendant sa réunion annuelle en septembre 1996 à Lisbonne (Portugal), la CENL décide de prendre Gabriel sous son ombrelle et de lancer un portail officiel plus conséquent à compter de janvier 1997.

Désormais trilingue (anglais, allemand, français), Gabriel est maintenu par la bibliothèque nationale des Pays-Bas (Koninklijke Bibliotheek), avec quatre sites miroir sur les serveurs des bibliothèques nationales du Royaume-Uni, de Finlande, d’Allemagne et de Slovénie.

Beaucoup plus tard, en été 2005, Gabriel fusionnera avec le site web de l’European Library (lancé par la CENL en janvier 2004) pour proposer un portail commun aux 43 bibliothèques nationales européennes. Europeana verra le jour trois ans après, en novembre 2008, en tant que bibliothèque numérique européenne.

Et les bibliothèques publiques?

Qu’en est-il des bibliothèques publiques? La bibliothèque publique d’Helsinki (Finlande) est la première bibliothèque à créer un site web en février 1994. Quatre ans plus tard, le document «Internet and the Library Sphere» publié en novembre 1998 par la Commission européenne évalue à 1.000 environ le nombre de bibliothèques publiques disposant d’un site web.

Ces bibliothèques sont réparties dans 26 pays. Les pays les plus représentés sont la Finlande (avec 247 bibliothèques), la Suède (132 bibliothèques), le Royaume-Uni (112 bibliothèques), le Danemark (107 bibliothèques), l'Allemagne (102 bibliothèques), les Pays-Bas (72 bibliothèques), la Lituanie (51 bibliothèques), l'Espagne (56 bibliothèques) et la Norvège (45 bibliothèques). La Russie a un site commun pour 26 bibliothèques publiques de recherche. Les pays nouvellement représentés sont la République tchèque (29 bibliothèques) et le Portugal (3 bibliothèques).

Les sites sont hétérogènes. Certains se contentent de mentionner l’adresse postale de la bibliothèque et ses heures d’ouverture, tandis que d’autres proposent toute une gamme de services, y compris un accès direct à leur catalogue en ligne.

@@@@@@@

5. Gallica, bibliothèque numérique de la Bibliothèque nationale de France (BnF)

Gallica est inauguré en octobre 1997 avec des textes et des images du 19e siècle francophone, «siècle de l'édition et de la presse moderne, siècle du roman mais aussi des grandes synthèses historiques et philosophiques, siècle scientifique et technique». Gallica élargit ensuite son champ d'action et devient rapidement l'une des grandes bibliothèques numériques mondiales.

***

Les débuts de Gallica

À l'époque, en 1997 donc, le serveur de Gallica stocke 2.500 livres numérisés en mode image complétés par les 250 livres numérisés en mode texte de la base Frantext de l'INaLF (Institut national de la langue française).

Classés par discipline, ces livres sont complétés par une chronologie du 19e siècle et des synthèses sur les grands courants en histoire, sciences politiques, droit, économie, littérature, philosophie, sciences et histoire des sciences.

Le site propose aussi un échantillon de la future iconothèque numérique, à savoir le fonds du photographe Eugène Atget, une sélection de documents sur l'écrivain Pierre Loti, une collection d'images de l'École nationale des ponts et chaussées -- ces images ayant trait aux grands travaux de la révolution industrielle en France -- et enfin un choix de livres illustrés de la bibliothèque du Musée de l'Homme.

Fin 1997, Gallica se considère moins comme une banque de documents numérisés que comme un «laboratoire dont l'objet est d'évaluer les conditions d'accès et de consultation à distance des documents numériques». Le but est d'expérimenter la navigation dans les collections, en permettant le libre parcours du chercheur ou du lecteur curieux.

La fin des années 1990

Début 1998, Gallica annonce 100.000 volumes et 300.000 images pour la fin 1999. Sur les 100.000 volumes prévus, qui représenteraient 30 millions de pages numérisées, plus du tiers concernerait le 19e siècle.

Quant aux 300.000 images fixes, la moitié viendrait des départements spécialisés de la BnF (Estampes et photographie, Manuscrits, Arts du spectacle, Monnaies et médailles, etc.), et

l'autre moitié de collections d'établissements publics (musées et bibliothèques, Documentation française, École nationale des ponts et chaussées, Institut Pasteur, Observatoire de Paris, etc.) ou privés (agences de presse dont Magnum, l'Agence France-Presse, Sygma, Rapho, etc.).

En mai 1998, la BnF revoit ses espérances à la baisse et modifie quelque peu ses orientations premières. Jérôme Strazzulla, journaliste au quotidien Le Figaro, explique dans un article du 3 juin 1998 que la BnF est passée «d'une espérance universaliste, encyclopédique, à la nécessité de choix éditoriaux pointus».

Dans le même article, le président de la BnF, Jean-Pierre Angremy, rapporte la décision du comité éditorial de Gallica: «Nous avons décidé d'abandonner l'idée d'un vaste corpus encyclopédique de cent mille livres, auquel on pourrait sans cesse reprocher des trous. Nous nous orientons aujourd'hui vers des corpus thématiques, aussi complets que possibles, mais plus restreints. (...) Nous cherchons à répondre, en priorité, aux demandes des chercheurs et des lecteurs.»

Les années 2000

Cinq ans plus tard, en 2003, Gallica rassemble 70.000 ouvrages et 80.000 images allant du Moyen-Âge au début du 20e siècle, tous documents libres de droits.

Mais la numérisation en mode image n'autorise pas la recherche textuelle alors que Gallica se trouve être la plus grande bibliothèque numérique francophone en nombre de titres disponibles en ligne.

La recherche textuelle est toutefois possible dans les tables des matières, les sommaires et les légendes des corpus iconographiques, qui sont numérisés en mode texte.

Seule une petite collection de livres (1.120 livres en février 2004) est intégralement numérisée en mode texte, celle de la base Frantext, intégrée à Gallica.

Tous problèmes auxquels la BnF remédie au fil des mois, avec une navigation plus aisée et la conversion progressive des livres du mode image au mode texte grâce à un logiciel OCR (Optical Character Recognition), avec possibilité donc de recherche textuelle.

En février 2005, Gallica compte 76.000 ouvrages.

À la même date, la BnF annonce la mise en ligne prochaine (entre 2006 et 2009) de la presse française parue entre 1826 et 1944, à savoir 22 titres représentant 3,5 millions de pages.

Début 2006, les premiers journaux disponibles en ligne sont les quotidiens Le Figaro (fondé en 1826), La Croix (fondée en 1883), L'Humanité (fondée en 1904) et Le Temps (fondé en 1861 et disparu en 1942).

En mars 2010, Gallica franchit la barre du million de documents -- livres, manuscrits, cartes, images, périodiques (presse et revues), fichiers sonores (paroles et musiques) et partitions musicales -- dont la plupart sont accessibles gratuitement sur un site dont l'interface désormais quadrilingue (français, anglais, espagnol, portugais) ne cesse de s'améliorer au fil des ans.

Si les documents sont logiquement en langue française dans leur très grande majorité, on y trouve aussi des documents en anglais, en italien, en allemand, en latin ou en grec selon les disciplines.

La numérisation en mode image

La numérisation en mode image consiste à scanner le livre, et correspond donc à la photographie du livre page après page. La présentation originale étant conservée, on peut feuilleter le livre à l'écran. La version informatique est le fac-similé numérique de la version imprimée.

C'est la méthode employée à la fin des années 1990 pour les programmes de numérisation à grande échelle, y compris pour Gallica. Ne sont numérisés en mode texte que les tables des matières, les sommaires et les corpus de documents iconographiques, afin de faciliter la recherche textuelle.

Pourquoi ne pas tout numériser en mode texte? La BnF répond en 2000 sur le site de Gallica: «Le mode image conserve l'aspect initial de l'original y compris ses éléments non textuels. Si le mode texte autorise des recherches riches et précises dans un document et permet une réduction significative du volume des fichiers manipulés, sa réalisation, soit par saisie soit par OCR, implique des coûts de traitement environ dix fois supérieurs à la simple numérisation. Ces techniques, parfaitement envisageables pour des volumes limités, ne pouvaient ici être économiquement justifiables au vu des 50.000 documents (représentant presque 15 millions de pages) mis en ligne.»

Dans les années qui suivent, Gallica convertira toutefois nombre de ses livres du mode image au mode texte pour permettre les recherches textuelles.

La numérisation en mode texte

La numérisation en mode texte consiste à scanner le livre en mode image, puis à le convertir en texte grâce à un logiciel OCR. La version texte ne conserve pas la présentation originale du livre ou de la page. Le livre devient un texte, à savoir un ensemble de caractères apparaissant en continu à l'écran.

Du fait du temps passé au traitement de chaque livre, ce mode de numérisation est assez long, et donc nettement plus coûteux que la numérisation en mode image. Dans de nombreux cas, il est toutefois préférable, puisqu'il permet l'indexation, la recherche textuelle, l'analyse textuelle, une étude comparative entre plusieurs textes ou plusieurs versions du même texte, etc.

C’est la méthode utilisée par exemple par le Projet Gutenberg, qui propose aujourd'hui la plus grande collection de livres numériques gratuits au format texte, avec des livres relus et corrigés à deux reprises par des milliers de volontaires pour être fiables à 99,95% par rapport à leur version imprimée.

L’utilité des deux modes de numérisation

Concepteur du logiciel Mot@mot, un logiciel de remise en page des fac-similés numériques, Pierre Schweitzer insiste sur l’utilité des deux modes de numérisation.

Il explique en janvier 2001 dans un entretien par courriel: «Le mode image permet d’avancer vite et à très faible coût. C’est important car la tâche de numérisation du domaine public est immense. Il faut tenir compte aussi des différentes éditions: la numérisation du patrimoine a pour but de faciliter l’accès aux œuvres, il serait paradoxal qu’elle aboutisse à se focaliser sur une édition et à abandonner l’accès aux autres.

Chacun des deux modes de numérisation s’applique de préférence à un type de document, ancien et fragile ou plus récent, libre de droit ou non (pour l’auteur ou pour l’édition), abondamment illustré ou pas.

Les deux modes ont aussi des statuts assez différents: en mode texte ça peut être une nouvelle édition d’une œuvre, en mode image c’est une sorte d’"édition d’édition", grâce à un de ses exemplaires (qui fonctionne alors comme une fonte d’imprimerie pour du papier). En pratique, le choix dépend bien sûr de la nature du fonds à numériser, des moyens et des buts à atteindre. Difficile de se passer d’une des deux façons de faire.»

@@@@@@@

 6. Du PDA au smartphone

En 2001, la liseuse n'a pas encore gagné son pari. On compte 17 millions de PDA (Personal Digital Assistants) dans le monde pour seulement 100.000 liseuses, d'après un Seybold Report publié en avril 2001. 13,2 millions de PDA sont vendus en 2001, dont le Palm Pilot (lancé en mars 1996) et le Pocket PC de Microsoft (lancé en mars 2000). En 2005, le PDA laisse progressivement la place au smartphone. L'iPhone d'Apple (lancé en juin 2007) devient le produit phare de toute une génération.

***

Le Palm Pilot

Basée en Californie, la société Palm lance en mars 1996 le Palm Pilot, premier PDA du marché, et vend 23 millions de Palm Pilot entre 1996 et 2002. Le système d'exploitation du Palm Pilot est le Palm OS et son logiciel de lecture est le Palm Reader, rejoint par le logiciel de lecture Mobipocket Reader en mars 2001.

En juillet 2002, le Palm Reader, jusque-là utilisable sur Palm Pilot et sur Pocket PC, est également utilisable sur ordinateur. À la même date, Palm Digital Media (renommée plus tard Palm eBook Store) distribue 5.500 titres dans plusieurs langues. Le catalogue comprend 10.000 titres un an plus tard.

Un changement majeur donc pour les adeptes du livre numérique qui, avant l'avènement du Palm Pilot, ne pouvaient lire les livres que sur l'écran de leur ordinateur, portable ou non.

Si certains professionnels du livre (et certains ophtalmologues) s'inquiètent de la petitesse de l'écran, les adeptes de la lecture sur PDA sont enthousiastes à l'idée de lire sur un appareil mobile multitâches et ne se plaignent guère de la taille de l'écran.

L'eBookMan (Franklin)

Lancé par la société Franklin en 2000, l'eBookMan est un PDA multimédia permettant de lire des livres numériques sur le logiciel de lecture Franklin Reader (tout comme le logiciel de lecture Mobipocket Reader en octobre 2001). L'eBookMan séduit les professionnels du livre technophiles puisqu'il reçoit l'eBook Technology Award de la Foire internationale de Francfort en octobre 2000.

Trois modèles (EBM-900, EBM-901 et EBM-911) sont disponibles début 2001, avec une mémoire vive de 8 ou 16 Mo selon les modèles et avec un écran LCD rétro-éclairé ou non. L'écran est nettement plus grand que celui des autres PDA du marché, mais n'existe qu'en noir et blanc, contrairement à la gamme Pocket PC de Microsoft ou à certains modèles du Palm Pilot avec écran couleur.

L'eBookMan permet l'écoute de livres audionumériques et de fichiers musicaux au format MP3. La société Franklin développe aussi une librairie numérique sur son site et passe des partenariats avec plusieurs sociétés, notamment avec Audible.com pour avoir accès à sa collection de 4.500 livres audionumériques.

Le logiciel de lecture Franklin Reader séduit puisqu'il est également proposé par défaut sur les gammes de PDA Psion, de Palm et de Pocket PC, tout comme sur le premier smartphone du marché lancé en 2001 par Nokia.

D'autres modèles de PDA

En 2002, la gamme Palm Pilot est toujours le leader du marché (avec 36,8% des PDA vendus), suivi de la gamme Pocket PC de Microsoft (lancée en mars 2000) et suivi aussi par les modèles de Hewlett-Packard, Sony, Handspring, Toshiba et Casio. Les systèmes d'exploitation utilisés sont essentiellement le Palm OS (pour 55% des PDA) et le Pocket PC (pour 25,7% des PDA).

Pour mémoire, les grands logiciels de lecture de l'époque sont le Mobipocket Reader (lancé en mars 2000), le Microsoft Reader (lancé en avril 2000), le Palm Reader (lancé en mars 2001), l'Acrobat Reader (lancé en mai 2001 pour le Palm Pilot et en décembre 2001 pour le Pocket PC) et l'Adobe Reader (lancé en mai 2003 pour remplacer l'Acrobat Reader).

En 2003, des centaines de nouveautés sont vendues en version numérique dans les eBookStores d'Amazon.com et Barnes & Noble.com et dans le Yahoo! eBookStore, et sur des sites d'éditeurs comme Random House ou PerfectBound.

À la même date, le catalogue de Palm Digital Media approche 10.000 titres disponibles pour PDA, avec 15 à 20 nouveaux titres par jour et 1.000 nouveaux clients par semaine. La société Mobipocket distribue 6.000 titres numériques dans plusieurs langues, soit sur son site soit dans des librairies partenaires. La librairie numérique Numilog distribue 3.500 titres numériques (livres et périodiques) en français et en anglais.

En 2004, les trois principaux fabricants de PDA sont Palm, Sony et Hewlett-Packard. Suivent Handspring, Toshiba, Casio et d'autres. Mais le PDA est de plus en plus concurrencé par le smartphone, qui est un téléphone portable doublé d'un PDA, et les ventes commencent à baisser. Signe des temps, Sony décide de se retirer du marché des PDA en février 2005.

Les smartphones

Le premier smartphone est le Nokia 9210, modèle précurseur lancé en 2001 par la société finlandaise Nokia, grand fabricant mondial de téléphones portables. La plateforme du Nokia 9210 est le Symbian OS. Suivent le Nokia Series 60, le Sony Ericsson P800, puis les modèles de Motorola et de Siemens. Ces différents modèles permettent de lire des livres numériques sur le logiciel de lecture Mobipocket Reader.

Appelé aussi téléphone multimédia, téléphone multifonctions ou encore téléphone intelligent, le smartphone dispose d'un écran couleur, du son polyphonique et de la fonction appareil photo, qui viennent s'ajouter aux fonctions habituelles du PDA (agenda, dictaphone, lecteur de livres numériques, lecteur de musique, etc.).

Les smartphones représentent 3,7% des ventes de téléphones mobiles en 2004 et 9% de ces ventes en 2006, à savoir 90 millions de smartphones pour un milliard de téléphones portables sur notre planète.

L'iPhone (Apple)

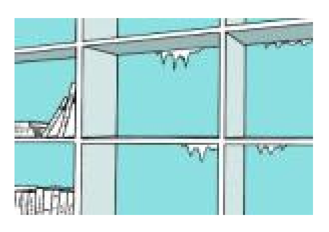
Présenté en janvier 2007 par Steve Jobs, l'iPhone, le smartphone d'Apple, est un téléphone mobile multifonctions qui intègre le baladeur de musique iPod (lancé en octobre 2001), un appareil photo et un navigateur web.

Le premier iPhone a les caractéristiques suivantes: grand écran tactile de 3,5 pouces, synchronisation automatique avec la plateforme iTunes pour télécharger musique et vidéos, appareil photo de 2 MP, navigateur Safari d'Apple, système d'exploitation Mac OS X, téléphonie par le biais des réseaux GSM et EDGE, connexion internet via la WiFi et enfin connexion Bluetooth. Qui dit mieux?

Par contre le prix n'est pas pour toutes les bourses. L'iPhone est lancé en juin 2007 aux États-Unis au prix de 499 dollars US pour le modèle de 4 Go et 599 dollars pour le modèle de 8 Go. Son lancement à l'international a lieu fin 2007 en Europe et en 2008 en Asie.

Signe des temps, Google Books lance en février 2009 un portail spécifique de livres numériques pour téléphone mobile et smartphone. Le catalogue comprend 1,5 million de titres du domaine public, auxquels s'ajoutent 500.000 autres titres téléchargeables hors des États-Unis, du fait d'une législation du copyright moins restrictive dans certains pays.

@@@@@@@



7. Les premières liseuses

Sans surprise, les premières liseuses voient le jour dans la Silicon Valley, terre des projets technologiques les plus fous, avec les fonds nécessaires pour les développer. Ces liseuses sont le Rocket eBook (de NuvoMedia) et le SoftBook Reader (de SoftBook Press), qui voient le jour en 1998. D'autres modèles suivent, par exemple le Gemstar eBook (de Gemstar) en 2000, le LIBRIe (de Sony) en 2004, le Cybook 2G (de Bookeen) à la même date, le Sony Reader (de Sony) en 2006, le Kindle (d'Amazon) en 2007 et le Nook (de Barnes & Noble) en 2009.

***

Les débuts de la liseuse

Alors que les liseuses étaient jusque-là l'apanage des films de science-fiction, les premières liseuses du marché -- le Rocket eBook de NuvoMedia et le SoftBook Reader de SoftBook Press -- suscitent un engouement certain en 1998, même si peu de gens vont jusqu'à les acheter.

Le problème n'est pas seulement leur prix prohibitif (plusieurs centaines de dollars), qui ne poserait guère de problème à certains dans la Silicon Valley. Le problème est surtout un choix de livres très restreint, le catalogue de livres numériques étant encore ridicule par rapport à la production imprimée. Les éditeurs commencent tout juste à produire des livres numériques et se demandent encore comment les commercialiser, la plupart étant tétanisés par les risques de piratage.

Ces deux liseuses fonctionnent sur batteries et disposent d'un écran à cristaux liquides (écran LCD: Liquid Cristal Display) noir et blanc rétro-éclairé ou non, avec une capacité de stockage d'une dizaine de livres. L'usager se connecte à l'internet soit par le biais d'un ordinateur (pour le Rocket eBook) soit directement grâce à un modem intégré (pour le SoftBook Reader) afin de télécharger des livres à partir des librairies numériques présentes sur les sites des sociétés NuvoMedia et SoftBook Press.

Le Rocket eBook (NuvoMedia)

Premier modèle du marché, le Rocket eBook est lancé en 1998 par NuvoMedia, une société californienne financée par la chaîne de librairies Barnes & Noble et le géant des médias Bertelsmann. NuvoMedia souhaite devenir «la solution pour distribuer des livres électroniques en procurant une infrastructure réseau aux éditeurs, distributeurs et usagers

afin de publier, distribuer, acheter et lire un contenu électronique de manière sécurisée et efficace sur l'internet».

La connexion entre le Rocket eBook et l'ordinateur (PC ou Macintosh) se fait par le biais du Rocket eBook Cradle, un périphérique à deux câbles, les deux cables étant d'une part un câble pour se connecter à une prise électrique par le biais d'un adaptateur et d'autre part un câble série pour se connecter à l'ordinateur.

Le SoftBook Reader (SoftBook Press)

Deuxième modèle du marché, le SoftBook Reader est lancé par SoftBook Press, une société californienne financée par les deux grandes maisons d'édition Random House et Simon & Schuster. Le SoftBook Reader utilise le SoftBook Network, «un service de distribution de contenu basé sur l'internet».

Cette liseuse permet aux lecteurs de «télécharger facilement, rapidement et de manière sécurisée un large choix de livres et de revues grâce à sa connexion internet intégrée. Contrairement à l'ordinateur, le SoftBook Reader possède une ergonomie conçue pour la lecture de longs documents et de livres.»

D'autres modèles suivent

D'autres modèles de liseuses sont lancés en 1999, toujours dans la Silicon Valley, par exemple l'EveryBook Reader de la société EveryBook et le Millennium eBook (le nouveau millénaire approche) de la société Librius.

L'EveryBook Reader est une liseuse à double écran lancée par EveryBook en tant que «bibliothèque vivante dans un simple livre» pouvant stocker 50 livres numériques, avec un modem intégré permettant l'accès à l'EveryBook Store, afin de «consulter, acheter et recevoir le texte intégral de livres, magazines et partitions de musique».

Le Millennium eBook est une liseuse «petite et bon marché» lancée par Librius, «une société de commerce électronique procurant un service complet». Sur le site de la société, un World Bookstore propose «des copies numériques de milliers de livres» disponibles via l'internet.

Toutes ces liseuses (Rocket eBook, SoftBook Reader, EveryBook Reader, Millennium eBook et d'autres) pèsent entre 700 grammes et deux kilos et ne durent pas. Il faudra attendre le nouveau millénaire pour voir apparaître des modèles ayant une durée de vie légèrement plus longue, par exemple le Gemstar eBook lancé en novembre 2000 aux États-Unis et le Cybook (première génération, celui de Cytale) lancé en janvier 2001 en Europe.

Le Gemstar eBook (Gemstar)

Gemstar est une société américaine spécialisée dans les produits et services numériques pour les médias et souhaite lancer sa propre liseuse. Gemstar lance le Gemstar eBook en novembre 2000 aux États-Unis après avoir racheté les sociétés californiennes Nuvomedia (auteure du Rocket eBook) et SoftBook Press (auteure du SoftBook Reader).

Le Gemstar eBook se décline en deux modèles -- le REB 1100 (écran noir et blanc, successeur du Rocket eBook) et le REB 1200 (écran couleur, successeur du SoftBook Reader) -- construits et vendus sous le label RCA, un label qui appartient à Thomson Multimedia. La plateforme, le navigateur et le logiciel de lecture sont spécifiques à l'appareil, tout comme le format de lecture, basé sur le format OeB (Open eBook, ancêtre de l'EPUB).

Les deux modèles sont vendus respectivement 300 et 699 dollars US par la chaîne de magasins SkyMall. Mais les ventes sont très inférieures aux pronostics. En avril 2002, un article du New York Times annonce l'arrêt de la fabrication de ces liseuses par RCA. C'est chose faite peu après.

Lancés à l'automne 2002, les modèles suivants -- le GEB 1150 et le GEB 2150 -- sont produits cette fois sous le label Gemstar et vendus par SkyMall à un prix beaucoup plus compétitif, avec ou sans abonnement annuel ou bisannuel à la librairie numérique du Gemstar eBook.

Le GEB 1150 coûte 199 dollars US sans abonnement, et 99 dollars avec abonnement annuel (facturé 20 dollars par mois). Le GEB 2150 coûte 349 dollars sans abonnement, et 199 dollars avec abonnement bisannuel (facturé lui aussi 20 dollars par mois).

Malgré tous ces efforts, les ventes restent peu concluantes -- faute d'un marché mûr pour ce genre d'appareil -- et Gemstar décide de mettre fin à ses activités eBook. La société cesse la vente de ses liseuses en juin 2003 et la vente de ses livres numériques le mois suivant.

Le Cybook (Cytale puis Bookeen)

Première liseuse européenne, le Cybook est lancé en janvier 2001 par la société française Cytale. Sa mémoire -- 32 Mo de mémoire SDRAM et 16 Mo de mémoire flash -- permet de stocker 15.000 pages de texte, soit 30 livres de 500 pages, dans un appareil de format moyen (21 x 16 cm) pesant un kilo.

Olivier Pujol, PDG de Cytale, explique en décembre 2000 dans un entretien par courriel: «J'ai croisé il y a deux ans le chemin balbutiant d'un projet extraordinaire, le livre électronique. Depuis ce jour, je suis devenu le promoteur impénitent de ce nouveau mode d'accès à l'écrit, à la lecture, et au bonheur de lire. La lecture numérique se développe enfin, grâce à cet objet merveilleux: bibliothèque, librairie nomade, livre "adaptable", et aussi moyen d'accès à tous les sites littéraires (ou non), et à toutes les nouvelles formes de la littérature, car c'est également une fenêtre sur le web.»

Mais les ventes du Cybook sont elles aussi très inférieures aux pronostics et forcent la société à cesser ses activités en juillet 2002. La commercialisation du Cybook est reprise par la société Bookeen, créée en 2003 à l'initiative de Michael Dahan et Laurent Picard, deux ingénieurs de Cytale. Le Cybook 2G (2e génération) est lancé en avril 2004 et se décline en plusieurs modèles. Suit le Cybook Gen3 (3e génération) en juillet 2007, avec un écran utilisant pour la première fois la technologie E Ink.

Le LIBRIe et le Sony Reader (Sony)

La société Sony lance en avril 2004 sa première liseuse, le LIBRIe 1000-EP, commercialisé au Japon et produit en partenariat avec les sociétés Philips et E Ink. Le LIBRIe est la première liseuse du marché à utiliser la technologie d'affichage développée par la société E Ink.

Cette liseuse pèse 300 grammes avec piles et protection d'écran (loin du poids des premières liseuses, donc) pour une taille de 12,6 x 19 x 1,3 centimètres. Elle fonctionne avec quatre piles alcalines. Sa mémoire est de 10 Mo (avec possibilité d'extension) et sa capacité de stockage de 500 livres. Son écran de 6 pouces a une définition de 170 DPI et une résolution de 800 x 600 pixels. Un port USB permet le téléchargement des livres à partir d'un ordinateur. La liseuse comprend aussi un clavier, une fonction d'enregistrement et une synthèse vocale.

Suite au LIBRIe lancé en avril 2004 au Japon, Sony lance le Sony Reader en octobre 2006 aux États-Unis. L'écran de cette liseuse utilise une technologie E Ink plus avancée. Selon Michael Cook, auteur du site epubBooks.com, il s'agit d'«un écran qui donne une excellente expérience de lecture, très proche de celle du vrai papier, et qui ne fatigue pas les yeux».

Un autre avantage de cette liseuse sur ses concurrentes est la durée de vie de la batterie, avec plus de 7.000 pages consultables, ou deux semaines sans nécessité de la recharger (cela reste à voir quand même). Cette liseuse est aussi la première à utiliser Adobe Digital Editions, un logiciel qui adapte le texte du livre à la taille de l'écran. Le Sony Reader est progressivement disponible au Canada, au Royaume-Uni, en Allemagne et en France.

Le Kindle (Amazon)

Sans surprise, la grande librairie en ligne Amazon.com lance en novembre 2007 sa propre liseuse, le Kindle, avec un format livresque (19 x 13 x 1,8 cm), un poids de 290 grammes, un écran noir et blanc de 6 pouces avec une résolution de 800 x 600 pixels, un clavier, une mémoire de 256 Mo (extensible par carte SD), un port USB et la possibilité de se connecter à l'internet via la WiFi. Le Kindle peut contenir jusqu'à 200 livres sur les 80.000 livres numériques que propose le catalogue d'Amazon, ce qui est un bon début.

Amazon lance ensuite en février 2009 le Kindle 2 pour un prix plus modique, qui continue de baisser sensiblement dans les mois qui suivent, puis le Kindle DX en mai 2009 avec un écran de 9,7 pouces permettant la lecture de journaux et magazines.

Le catalogue d'Amazon comptabiliserait 450.000 titres numériques en mars 2010, y compris des livres et revues audionumériques suite au rachat du catalogue d'Audible.com en janvier 2009.

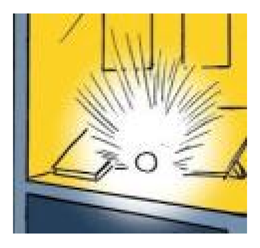
Le Nook (Barnes & Noble)

En novembre 2009, la grande chaîne de librairies américaine Barnes & Noble lance sa propre liseuse, le Nook. Le Nook dispose d’une plateforme Android et d'un écran E Ink de 6 pouces, avec une connexion WiFi et une connexion 3G. Un modèle plus économique disposant de la seule connexion WiFi est lancé en juin 2010. Le Nook Color apparaît en octobre 2010 avec un écran LCD de 7 pouces pour la lecture de magazines et de livres d'images. Le catalogue de Barnes & Noble propose 2 millions de livres numériques à la fin 2010. Un nouveau Nook plus léger est lancé en mai 2011 sous plateforme Android, avec un écran de 6 pouces utilisant la technologie E Ink Pearl tactile.

***

Les tablettes multifonctions permettent elles aussi de lire des livres. L'iPad est lancé par Apple le 3 avril 2010 aux États-Unis, avec un iBookstore de 60.000 livres numériques qui s'étoffe rapidement. Un lancement mondial suit en juin 2010. Après l'iPod (lancé en octobre 2001) puis l'iPhone (lancé en juin 2007), deux objets cultes auprès de toute une génération, Apple devient lui aussi un acteur de poids dans le marché du livre numérique. Apple lance l’iPad 2 en mars 2011 aux États-Unis, avec un lancement deux semaines plus tard dans d’autres pays.

@@@@@@@



8. E Ink, une technologie d'encre électronique

Les recherches sur l'encre électronique débutent à la fin des années 1990 au sein de la société E Ink pour offrir un meilleur écran aux liseuses en attendant le papier électronique. E Ink présente le prototype de son écran en 2002 et le commercialise en 2004. Un projet un peu différent est la technologie Gyricon développée par Xerox. Radicalement novateur lui aussi dans un domaine connexe, le projet @folio voit le jour en France dès 1996 mais reste malheureusement à l'état de prototype faute de soutien financier.

***

La technologie E Ink

Des chercheurs du Media Lab du MIT (Massachusetts Institute of Technology) créent la société E Ink en avril 1997 pour développer une technologie d'encre électronique, d'abord pour les écrans rigides des liseuses, puis pour leurs écrans souples et enfin pour le futur papier électronique. Si le concept est révolutionnaire, cette technologie est le résultat d'une fusion entre trois sciences: la chimie, la physique et l'électronique.

Très schématiquement, la technologie utilisée par E Ink est la suivante: prises entre deux feuilles de plastique souple, des millions de micro-capsules contiennent chacune des particules noires et blanches en suspension dans un fluide clair. Un champ électrique positif ou négatif permet de faire apparaître le groupe de particules souhaité à la surface du support, pour afficher, modifier ou effacer les données.

En juillet 2002, E Ink présente le prototype du premier écran utilisant cette technologie, un écran de haute résolution à matrice active développé en partenariat avec les sociétés Toppan et Philips. Cet écran est commercialisé en 2004 pour diverses liseuses (LIBRIe, Sony Reader, Cybook, Kindle, Nook, etc.).

La première liseuse disposant d'un écran E Ink de 6 pouces (au lieu de l'écran à cristaux liquides habituel) est le LIBRIe, lancé en avril 2004 par Sony au Japon. Suit le Sony Reader lancé en octobre 2006 aux États-Unis, avec un écran utilisant une technologie E Ink plus avancée, à savoir «un écran qui donne une excellente expérience de lecture, très proche de celle du vrai papier, et qui ne fatigue pas les yeux», selon Michael Cook, auteur du site epubBooks.com.

Le CyBook Gen3 (lancé par Bookeen en juillet 2007), le Kindle (lancé par Amazon en novembre 2007) et le Nook (lancé par Barnes & Noble en novembre 2009) disposent eux aussi d'un écran E Ink.

La technologie Gyricon

Un autre projet d’encre électronique est développé par la société Xerox. PARC (Palo Alto Research Center), son centre de recherche dans la Silicon Valley, met au point une technique d’affichage un peu différente dénommée Gyricon.

Très schématiquement, la technologie est la suivante: prises entre deux feuilles de plastique souple, des millions de micro-alvéoles contiennent des micro-billes bicolores en suspension dans un liquide clair. Chaque bille est pourvue d'une charge électrique. Une impulsion électrique extérieure permet la rotation des billes, et donc le changement de couleur, pour afficher, modifier ou effacer des données.

Des chercheurs de PARC créent en décembre 2000 la société Gyricon Media dans le but de développer et commercialiser le SmartPaper, un papier électronique destiné à être produit en rouleaux, tout comme le papier traditionnel. Le marché pressenti est d'abord celui de l'affichage commercial, qui utilise le système SmartSign, développé en complément du SmartPaper.

La vente d’affichettes fonctionnant sur piles débute en 2004. Viennent ensuite les panneaux de signalisation électroniques puis les premiers prototypes de papier électronique et de journal électronique. Mais il est encore un peu tôt pour ce nouveau marché. Gyricon Media ferme ses portes en 2005 et les activités de recherche et de développement reviennent dans leur lieu premier à Xerox.

Le projet @folio

Une digression maintenant, pour rendre hommage à Pierre Schweitzer, architecte designer à Strasbourg (France), qui conçoit dès 1996 @folio, un baladeur de textes radicalement novateur et malheureusement resté à l’état de prototype, malgré le dépôt d’un brevet international en avril 2001 et la création de la start-up française iCodex en juillet 2002.

Pierre Schweitzer explique en août 2007 dans un entretien par courriel: «Il s’agit d’offrir un support de lecture efficace aux textes qui n’en ont pas, ceux qui sont accessibles sur le web. Avec @folio, je reste persuadé qu’un support de lecture transportable qui serait à la fois simple et léger, annotable et effaçable, à bas coût, respectueux de la page et de nos traditions typographiques, pourrait apporter un supplément de confort appréciable à tous les usagers du texte numérique. Une ardoise dont on pourrait feuilleter l’hypertexte à main nue, en lieu et place de l’imprimante. (...)

C'est un mode de représentation totalement ouvert et accessible à tous: il supporte l'écriture manuscrite, la calligraphie, les écritures non alphabétiques, et le dessin à main levée, toutes choses qui sont très difficiles à faire à l'aide d'un seul outil sur un ordinateur ou un "ebook" classique. Cette conception technique nouvelle et très simplifiée permet de recueillir une grande variété de contenus et surtout, elle permet un prix de vente très raisonnable (100 euros pour le modèle de base) dans différentes combinaisons de formats (tailles d'écran) et de mémoire (nombre de pages) adaptées aux différentes pratiques de lecture.»

Mais @folio restera un prototype et ne sera malheureusement jamais commercialisé faute de soutien financier en France et en Europe. Mon sentiment est qu'@folio aurait pu figurer en bonne place à côté du Kindle et de l'iPad, mais avec un modèle ouvert radicalement différent. L'exception française, qui ne s'est pas produite. Il ne reste qu'un beau souvenir immortalisé dans un article d'ActuaLitté daté de juin 2011.

@@@@@@@



9. Des dictionnaires et encyclopédies en ligne

Disponibles en ligne au tournant du millénaire, les premiers dictionnaires et encyclopédies de référence émanent de versions imprimées, par exemple le Dictionnaire universel francophone (Hachette), WebEncyclo (Atlas) ou encore le site de l'Encyclopaedia Universalis. Suivent des outils de référence conçus directement pour le web, par exemple le Grand dictionnaire terminologique (GDT) du Québec et WordReference.com. Wikipédia fera l'objet d'un chapitre spécifique.

***

Les premiers dictionnaires en ligne

Disponible en ligne dès 1997, le premier grand dictionnaire de langue française en accès libre et gratuit est le Dictionnaire universel francophone en ligne, qui répertorie 45.000 mots et 116.000 définitions tout en présentant «sur un pied d'égalité le français dit "standard" et les mots et expressions en français tel qu'on le parle sur les cinq continents». Issu de la collaboration entre Hachette et l'AUPELF-UREF (devenu depuis l'AUF -- Agence universitaire de la Francophonie), il correspond à la partie Noms communs du dictionnaire imprimé disponible chez Hachette.

L'équivalent pour la langue anglaise est le site Merriam-Webster OnLine, qui donne librement accès au Collegiate Dictionary, au Collegiate Thesaurus et à d'autres outils linguistiques.

En mars 2000, les 20 volumes de l'Oxford English Dictionary (OED) sont mis en ligne par l'Oxford University Press (OUP). La consultation du site est payante. Le site offre une mise à jour trimestrielle avec près de 1.000 entrées nouvelles ou révisées.

En mars 2002, deux ans après cette première expérience, l'Oxford University Press (OUP) lance l'Oxford Reference Online (ORO), une vaste encyclopédie conçue cette fois directement pour le web et consultable elle aussi sur abonnement payant. Avec 60.000 pages et un million d'entrées, elle représente l'équivalent d'une centaine d'ouvrages de référence.

Les premières encyclopédies en ligne

Mis en ligne en décembre 1999 par les éditions Atlas, WebEncyclo est la première grande encyclopédie francophone en accès libre et gratuit. La recherche est possible par mots-clés, par thèmes, par médias (cartes, liens internet, photos, illustrations) et par idées. Un appel à contribution incite les spécialistes d'un sujet donné à envoyer des articles, qui sont regroupés dans la section WebEncyclo contributif. Après avoir été libre, l'accès est ensuite soumis à une inscription préalable gratuite.

La mise en ligne de l'Encyclopaedia Universalis à la même date représente un ensemble de 28.000 articles signés de 4.000 auteurs. Si la consultation est payante sur la base d'un abonnement annuel, de nombreux articles sont également en accès libre.

Mis en ligne lui aussi en décembre 1999, le site Britannica.com est la première grande encyclopédie anglophone en accès libre et gratuit. Le site propose l'équivalent numérique des 32 volumes de l'Encyclopaedia Britannica (15e édition), en complément de la version imprimée et de la version CD-Rom, toutes deux payantes.

Britannica.com ajoute une sélection d'articles issus de 70 magazines, un guide des meilleurs sites, un choix de livres, etc., avec un moteur de recherche unique. En septembre 2000, le site fait partie des cent sites les plus visités du web. En juillet 2001, la consultation devient payante sur la base d'un abonnement annuel ou mensuel. Beaucoup plus tard, en 2009, Britannica.com ouvre son site à des contributeurs externes pour écrire et modifier des articles.

Le Grand dictionnaire terminologique (GDT)

Conçu lui aussi directement pour le web, avec accès libre et gratuit, le Grand dictionnaire terminologique (GDT) est un dictionnaire bilingue français-anglais de 3 millions de termes appartenant au vocabulaire industriel, scientifique et commercial. Sa mise en ligne en septembre 2000 est le résultat d'un partenariat entre l'Office québécois de la langue française (OQLF), auteur du dictionnaire, et Semantix, société spécialisée dans les solutions logicielles linguistiques.

Dès le premier mois, le GDT est consulté par 1,3 million de personnes, avec 60.000 requêtes par jour. La gestion du GDT est ensuite assurée par Convera Canada, avec 3,5 millions de requêtes mensuelles en février 2003. Une nouvelle version du GDT est mise en ligne en mars 2003. La gestion du dictionnaire est désormais assurée par l'OQLF lui-même, et non plus par une société prestataire, avec l’ajout du latin comme troisième langue.

WordReference.com

Michael Kellogg crée le site WordReference.com en 1999 pour proposer des dictionnaires bilingues gratuits en ligne, tout comme d’autres outils linguistiques et des forums de langues, avec l’ajout de nouvelles fonctionnalités au fil des ans.

Il raconte sur le site: «Mon site s’est progressivement développé pour devenir l’un des sites de dictionnaires en ligne les plus utilisés, et le principal dictionnaire en ligne pour les paires de langues anglais-espagnol, anglais-français, anglais-italien, espagnol-français et espagnol-portugais. Mon site est toujours classé sans interruption parmi les 500 sites les plus visités du web.»

Dans les années qui suivent, WordReference.com propose un dictionnaire anglais monolingue et des dictionnaires de l’anglais vers d’autres langues (arabe, chinois, coréen, français, grec, italien, japonais, polonais, portugais, roumain, tchèque, turc) et vice versa. Pour l’espagnol, on trouve un dictionnaire monolingue, un dictionnaire de synonymes, un dictionnaire espagnol-français et un dictionnaire espagnol-portugais. Des tableaux de conjugaison sont disponibles pour l’espagnol, le français et l’italien.

Lancé en 2010, WordReference Mini est une version miniature du site qui permet son intégration dans d’autres sites, par exemple des sites d’apprentissage de langues.

Les dictionnaires les plus consultés (anglais-espagnol, anglais-français, anglais-italien et vice versa) ont aussi leur version pour appareil mobile, avec d’autres paires de langues à venir.

@@@@@@@



10. Les aventures d'auteurs de best-sellers

Le maître du suspense américain Stephen King est le premier auteur de best-sellers à se lancer dans l'aventure numérique en 2000 malgré les risques commerciaux encourus. Il auto-publie un roman épistolaire inédit en plusieurs épisodes sur un site web créé a cet effet, avec un résultat mitigé, avant de nouvelles expériences avec son éditeur. D'autres auteurs de best-sellers tentent eux aussi l'aventure numérique avec leurs éditeurs, par exemple Frederick Forsyth en Grande-Bretagne, Arturo Pérez-Reverte en Espagne et Paulo Coelho au Brésil.

***

Stephen King prend des risques

Stephen King tâte d'abord le terrain numérique en mars 2000 en distribuant uniquement sur l'internet sa nouvelle «Riding the Bullet», assez volumineuse puisqu'elle comprend 66 pages. Du fait de la notoriété de l'auteur et de la couverture médiatique de ce scoop, cette expérience est un succès immédiat, avec 400.000 exemplaires téléchargés en 24 heures dans les librairies en ligne qui la vendent au prix de 2,5 dollars US.

Fort de cette expérience prometteuse, Stephen King décide de se passer des services de Simon & Schuster, son éditeur habituel. Il débute en juillet 2000 l'auto-publication en épisodes de son roman épistolaire inédit «The Plant» sur un site web dédié.

«The Plant» raconte l'histoire d'une plante carnivore s'emparant d'une maison d'édition et lui promettant le succès commercial en échange de sacrifices humains. Le premier chapitre est téléchargeable dans plusieurs formats (PDF, OeB, HTML, TXT) pour la modeste somme de 1 dollar US, avec paiement différé ou paiement immédiat sur le site d'Amazon.

«Le pire cauchemar des éditeurs»

Dans une lettre aux lecteurs publiée sur son site, Stephen King raconte que la création du site, le design et la publicité lui ont coûté la modeste somme de 124.150 dollars (!), sans compter sa prestation en tant qu'écrivain ni la rémunération de son assistante. Il précise aussi que la publication des chapitres suivants est liée au paiement du premier chapitre par au moins 75% des internautes.

«Mes amis, vous avez l’occasion de devenir le pire cauchemar des éditeurs», déclare-t-il dans sa lettre. «Comme vous le voyez, c’est simple. Pas de cryptage assommant! Vous voulez imprimer l’histoire et en faire profiter un(e) ami(e)? Allez-y. Une seule condition: tout repose sur la confiance, tout simplement. C’est la seule solution. Je compte sur deux facteurs. Le premier est l’honnêteté. Prenez ce que bon vous semble et payez pour cela, dit le proverbe. Le second est que vous aimerez suffisamment l’histoire pour vouloir en lire davantage. Si vous le souhaitez vraiment, vous devez payer. Rappelez-vous: payez, et l’histoire continue; volez, et l’histoire s’arrête.»

Une semaine après la mise en ligne du premier chapitre, on compte 152.132 téléchargements, avec paiement par 76% des lecteurs. Certains paient bien jusqu’à 10 ou 20 dollars (soit bien davantage que le dollar requis) pour compenser le manque à gagner de ceux qui ne paieraient pas, et éviter ainsi que la série ne s’arrête. La barre des 75% de lecteurs qui paient est dépassée de peu, au grand soulagement des fans, si bien que le deuxième chapitre suit un mois après.

La suite de l’expérience

En août 2000, dans une nouvelle lettre aux lecteurs, Stephen King annonce un nombre de téléchargements légèrement inférieur à celui du premier chapitre. Il en attribue la cause à une publicité moindre et à des problèmes de téléchargement. Mais le nombre de paiements est en nette diminution, les internautes ne réglant leur dû qu’une seule fois pour plusieurs téléchargements.

L’auteur s’engage toutefois à publier le troisième chapitre fin septembre, à la date prévue, et à prendre une décision ensuite sur la poursuite ou non de l’expérience, en fonction du nombre de paiements. Il annonce 11 ou 12 chapitres en tout, tout comme la gratuité du dernier chapitre, et il espère un nombre total de 1,7 million de téléchargements.

Plus volumineux (10.000 signes au lieu de 5.000 signes), les chapitres 4 et 5 passent à 2 dollars. Mais le nombre de téléchargements et de paiements ne cesse de décliner, avec 40.000 téléchargements seulement pour le cinquième chapitre alors que le premier chapitre avait été téléchargé 120.000 fois, et paiement pour 46% des téléchargements seulement.

L’arrêt de l’expérience

Fin novembre 2000, Stephen King annonce l’interruption de la publication de «The Plant» pendant une période indéterminée, après la parution du sixième chapitre, téléchargeable gratuitement à la mi-décembre. «"The Plant" va retourner en hibernation afin que je puisse continuer à travailler», précise-t-il sur son site. «Mes agents insistent sur la nécessité d’observer une pause afin que la traduction et la publication à l’étranger puissent aller de pair avec la publication en anglais.» Mais cette décision semble d’abord liée à l’échec commercial de l’expérience.

Cet arrêt suscite les vives critiques de plusieurs journalistes et critiques littéraires affirmant que Stephen King se ridiculise aux yeux du monde entier, preuve qu’ils n’ont visiblement pas

suivi l'histoire depuis ses débuts. L'auteur avait d'emblée annoncé la couleur puisqu'il avait lié la poursuite de la publication à un pourcentage de paiements satisfaisant.

Ils oublient aussi de reconnaître à Stephen King au moins un mérite, celui d'avoir été le premier à se lancer dans l'aventure, avec les risques qu'elle comporte. Entre juillet et décembre 2000, pendant les six mois qu'elle aura duré, nombreux sont ceux qui suivent les tribulations de «The Plant», à commencer par les éditeurs, quelque peu inquiets face à ce modèle économique inédit qui pourrait un jour concurrencer le circuit traditionnel.

Les éditeurs restent utiles

Qu'est-il advenu ensuite des expériences numériques de Stephen King? L'auteur reste très présent dans ce domaine, mais cette fois par le biais de son éditeur.

En mars 2001, son roman «Dreamcatcher» est le premier roman à être lancé simultanément en version imprimée par Simon & Schuster et en version numérique par Palm Digital Media, la librairie numérique de Palm, pour lecture sur les PDA Palm Pilot et Pocket PC.

En mars 2002, son recueil de nouvelles «Everything's Eventual» est lui aussi publié simultanément en deux versions: en version imprimée par Scribner, subdivision de Simon & Schuster, et en version numérique par Palm Digital Media, qui en propose un extrait en téléchargement libre. Et ainsi de suite, preuve que les éditeurs sont toujours utiles.

Stephen King ouvre la voie à d'autres auteurs de best-sellers pour leurs propres expériences numériques, par exemple Frederick Forsyth, maître britannique du thriller, Arturo Pérez-Reverte, romancier espagnol, et Paulo Coelho, romancier brésilien. Mais, forts de l'expérience d'auto-publication de Stephen King peut- être, ils n'ont pas l'intention de se passer de leurs éditeurs.

Frederick Forsyth

Frederick Forsyth, maître britannique du thriller, aborde la publication numérique avec l'appui de l'éditeur électronique londonien Online Originals, qui publie en novembre 2000 «The Veteran», la première nouvelle de «Quintet», une série de cinq nouvelles annoncées dans l'ordre suivant: «The Veteran», «The Miracle», «The Citizen», «The Art of the Matter» et «Draco».

Cette première nouvelle est vendue au prix de 3,99 pounds (6,60 euros) sur le site de l'éditeur et dans plusieurs librairies en ligne au Royaume-Uni (Alphabetstreet, BOL.com, WHSmith) et aux États-Unis (Barnes & Noble, Contentville, Glassbook).

Frederick Forsyth déclare à la même date sur le site d'Online Originals: «La publication en ligne sera essentielle à l'avenir. Elle crée un lien simple et surtout rapide et direct entre le producteur original (l'auteur) et le consommateur final (le lecteur), avec très peu d'intermédiaires. Il est passionnant de participer à cette expérience. Je ne suis absolument pas un spécialiste des nouvelles technologies. Je n'ai jamais vu de liseuse. Mais je n'ai

jamais vu non plus de moteur de Formule 1, ce qui ne m'empêche pas de constater combien ces voitures de course sont rapides.»

Toutefois cette première expérience ne dure pas, les ventes étant très inférieures aux pronostics. «Quintet» poursuivra une belle carrière imprimée, en attendant des jours meilleurs pour sa version numérique.

Arturo Pérez-Reverte

La première expérience numérique d'Arturo Pérez-Reverte est un peu différente. La série best-seller du romancier espagnol relate les aventures du Capitan Alatriste au 17e siècle. Le nouveau titre à paraître fin 2000 s'intitule «El oro del rey» (L'or du roi), bien avant le film Alatriste (2006), une superproduction espagnole de 20 millions d'euros.

En novembre 2000, en collaboration avec son éditeur Alfaguara, l'auteur décide de diffuser ce nouveau titre en version numérique sur une page spécifique du portail Inicia, en exclusivité pendant un mois, avant sa sortie en librairie. Le roman est disponible au format PDF pour 2,90 euros, un prix très inférieur aux 15,10 euros annoncés pour le livre imprimé.

Résultat de l'expérience, le nombre de téléchargements est très satisfaisant, mais pas celui des paiements. Un mois après la mise en ligne du roman, on compte 332.000 téléchargements, avec paiement par 12.000 lecteurs seulement.

À la même date, Marilo Ruiz de Elvira, directrice de contenus du portail Inicia, explique dans un communiqué: «Pour tout acheteur du livre numérique, il y avait une clé pour le télécharger sous 48 heures sur le site internet. Surtout au début, de nombreux internautes se sont échangés ce code d'accès dans les forums et ils ont donc téléchargé leur exemplaire sans payer. On a voulu faire un test et cela faisait partie du jeu. Arturo Pérez-Reverte voulait surtout qu'on le lise.»

Paulo Coelho

Paulo Coelho, romancier brésilien, est mondialement connu pour son roman «L'Alchimiste». En 2003, ses livres, traduits en 56 langues, ont été vendus en 53 millions d'exemplaires dans 155 pays, dont 6,5 millions d'exemplaires dans les pays francophones.

En mars 2003, Paulo Coelho décide de distribuer plusieurs de ses romans gratuitement en version PDF, en diverses langues, avec l'accord de ses éditeurs respectifs, dont Anne Carrière, son éditrice en France. Trois romans sont disponibles en français: «Manuel du guerrier de la lumière», «La cinquième montagne» et «Veronika décide de mourir».

Pourquoi une telle décision? L'auteur déclare à la même date par le biais de son éditrice: «Comme le français est présent, à plus ou moins grande échelle, dans le monde entier, je recevais sans cesse des courriers électroniques d'universités et de personnes habitant loin de la France, qui ne trouvaient pas mes œuvres.»

À la question classique relative au manque à gagner sur les ventes futures, Paulo Coelho répond: «Seule une minorité de gens a accès à l'internet, et le livre au format ebook ne remplacera jamais le livre papier.» Une remarque très juste en 2003, mais qui n'est plus de mise quelques années plus tard. Paulo Coelho réitère toutefois l’expérience avec d’autres titres au printemps 2011, pour la plus grande joie de ses lecteurs.

@@@@@@@



11. L’EPUB devient le format standard

Le format EPUB succède au format OeB (Open eBook) en avril 2005 en tant que standard du livre numérique et supplante peu à peu le format pionnier PDF. «Totalement différent du PDF, le format EPUB permet d’afficher un texte ajusté au tout dernier moment en fonction de l’écran et des choix du lecteur au moment précis de la lecture. Plusieurs applications permettent de créer un fichier EPUB directement, dont des logiciels libres. La plupart des liseuses sont compatibles avec le format EPUB.» (Nicolas Pewny, consultant en édition électronique)

***

La prolifération des formats

Les années 1998-2001 sont marquées par la prolifération des formats, chacun lançant son propre format de livre numérique dans le cadre d’un marché naissant promis à une expansion rapide.

Aux formats classiques -- formats TXT (texte), DOC (Microsoft Word), HTML (HyperText Markup Language), XML (eXtensible Markup Language) et PDF (Portable Document Format) -- s’ajoutent des formats propriétaires créés par plusieurs sociétés pour leurs propres logiciels de lecture. Ces logiciels correspondent souvent à un appareil donné (PDA, smartphone ou liseuse) et ne peuvent pas être utilisés sur d'autres appareils.

Au tournant du millénaire, nous avons par exemple le Palm Reader (pour le Palm Pilot), le Rocket eBook Reader (pour le Rocket eBook), le Franklin Reader (pour l'eBookMan), le logiciel de lecture Cytale (pour le Cybook), le Gemstar eBook Reader (pour le Gemstar eBook), sans compter des logiciels comme le Glassbook Reader (pour les livres de la société Glassbook, rachetée par Palm) ou le Peanut Reader (pour les livres de la société Peanut Press, rachetée par Adobe).

Le format OeB (Open eBook)

À l’instigation du National Institute of Standards & Technology (NIST) aux États-Unis, l’Open eBook Initiative voit le jour en juin 1998 et forme un groupe de travail de 25 personnes sous le nom d'Open eBook Authoring Group. Ce groupe élabore l’OeB (Open eBook), un format de livre numérique basé sur le langage XML et destiné à normaliser le contenu, la structure et la présentation des livres numériques.

Le format OeB est défini par l'OeBPS (Open eBook Publication Structure), dont la version 1.0 est disponible en septembre 1999. Téléchargeable gratuitement, l'OeBPS dispose d'une version ouverte et gratuite appartenant au domaine public. La version originale est destinée aux professionnels de la publication puisqu'elle doit souvent être associée à une technologie normalisée de gestion des droits numériques, et donc à un système de DRM (Digital Rights Management) permettant de contrôler l'accès des livres numériques sous droits.

Fondé en janvier 2000 pour prendre la suite de l'Open eBook Initiative, l'Open eBook Forum (OeBF) est un consortium industriel international regroupant constructeurs, concepteurs de logiciels, éditeurs, libraires et spécialistes du numérique (avec 85 participants en 2002) dans l'optique de développer le format OeB et l'OeBPS.

Le format OeB devient un standard qui sert lui-même de base à de nombreux formats, par exemple le format LIT pour le Microsoft Reader ou le format PRC pour le Mobipocket Reader.

Le format EPUB

En avril 2005, l'Open eBook Forum devient l'International Digital Publishing Forum (IDPF), et le format OeB laisse la place au format EPUB, acronyme de «Electronic PUBlication». (Douze ans plus tard, en janvier 2017, l'IDPF est intégrée au World Wide Web Consortium (W3C) qui poursuit le développement de l'EPUB.)

Le format EPUB est rapidement utilisé par les éditeurs parce qu'il facilite la mise en page des livres sur tout appareil de lecture (ordinateur, PDA, smartphone, liseuse) en fonction de la taille de l'écran. Les fichiers PDF (format pionnier du livre numérique) créés avec des versions récentes du logiciel Adobe Acrobat sont compatibles avec le format EPUB.

L'EPUB supplante peu à peu le PDF, avec une deuxième version (EPUB2) en mai 2010 et une troisième version (EPUB3) en octobre 2011.

Nicolas Pewny, consultant en édition électronique, précise en octobre 2015 dans un entretien par courriel: «Les PDF peuvent contenir des liens et des boutons, des champs de formulaire, des contenus audio et vidéo. Ils prennent également en charge les signatures électroniques. Le principal avantage de ce format est d'être parfaitement fidèle à la mise en page d'origine. Par contre, il est assez lourd et manque de souplesse; il ne s'adapte que très peu à l'écran utilisé.

Totalement différent du PDF, le format EPUB permet d'afficher un texte ajusté au tout dernier moment en fonction de l'écran et des choix du lecteur au moment précis de la lecture. La dernière version standardisée, EPUB3, repose sur HTML5, la dernière révision majeure d'HTML, format de données conçu pour représenter les pages web. Le format EPUB3 peut interroger les fonctionnalités du lecteur via JavaScript. Il peut ainsi adapter le contenu avec une grande souplesse.

À mon avis, EPUB3 est le meilleur format pour la publication électronique aujourd'hui. Plusieurs applications permettent de créer un fichier EPUB directement, dont des logiciels

libres. La plupart des liseuses sont compatibles avec le format EPUB, excepté le Kindle d’Amazon, qui nécessite une transformation (par exemple avec le logiciel libre Calibre). Diverses applications de lecture supportant l’EPUB sont également disponibles pour les ordinateurs personnels, les tablettes tactiles et les smartphones.»

@@@@@@@

12. Wikipédia, une encyclopédie planétaire

Lancée en janvier 2001 par Jimmy Wales et Larry Sanger, Wikipédia est une encyclopédie gratuite en ligne écrite collectivement et dont le contenu est librement réutilisable. Rédigés par des milliers de volontaires dans des dizaines de langues, les articles et les illustrations restent la propriété de leurs auteurs. Leur libre utilisation est régie par la licence GFDL puis par une licence Creative Commons. Wikipédia comprend 17 millions d'articles dans 270 langues en 2011, avec 400 millions de visiteurs par mois. D'autres encyclopédies collaboratives suivent, par exemple l'Encyclopedia of Life.

***

Les débuts de Wikipédia

Fondée en janvier 2001 à l'initiative de Jimmy Wales et Larry Sanger (Larry Sanger quitte plus tard l'équipe pour lancer Citizendium, une autre encyclopédie), Wikipédia est une encyclopédie gratuite en ligne écrite collectivement et dont le contenu est librement réutilisable.

Wikipédia est immédiatement très populaire. Sans publicité et financée par des dons, elle est rédigée par des milliers de volontaires dans des dizaines de langues, avec possibilité pour tout un chacun d'écrire, corriger et compléter les articles, aussi bien les siens que ceux d'autres contributeurs.

Les articles restent la propriété de leurs auteurs et leur libre utilisation est d'abord régie par la licence GFDL (GNU Free Documentation License) puis par une licence Creative Commons BY-SA. Les articles et les illustrations (images, photos, dessins, cartes, graphiques, etc.) peuvent donc être réutilisés par tous dans la mesure où la même licence est utilisée.

Fondée en juin 2003, la Wikimedia Foundation gère non seulement Wikipédia mais aussi Wiktionary (dictionnaire et thésaurus multilingue) créé en décembre 2002, Wikibooks (livres et manuels en cours de rédaction) créé en juin 2003, auxquels s'ajoutent ensuite Wikiquote (répertoire de citations), Wikisource (textes appartenant au domaine public), Wikimedia Commons (sources multimédia), Wikispecies (répertoire d'espèces animales et végétales), Wikinews (site d'actualités) et Wikiversity (matériel d'enseignement).

Quelques chiffres

En décembre 2004, Wikipédia compte 1,3 million d’articles rédigés par 13.000 contributeurs dans une centaine de langues. En décembre 2006, Wikipédia compte 6 millions d’articles dans 250 langues et devient l'un des dix sites les plus visités du web. En mai 2007, 7 millions
d'articles sont disponibles dans 192 langues, dont 1,8 million d’articles en anglais, 589.000 articles en allemand, 500.000 articles en français, 260.000 articles en portugais et 236.000 articles en espagnol.

En 2009, l’encyclopédie est l’un des cinq sites les plus visités du web. En septembre 2010, Wikipédia compte 14 millions d'articles dans 272 langues, dont 3,4 millions d’articles en anglais, 1,1 million d’articles en allemand et 1 million d’articles en français, qui est toujours la troisième langue de l'encyclopédie. Lors de ses dix ans en janvier 2011, Wikipédia compte 17 millions d’articles dans 270 langues et 400 millions de visiteurs par mois pour l’ensemble de ses sites.

Larry Sanger, co-fondateur de Wikipédia avec Jimmy Wales, fonde en mars 2007 une nouvelle encyclopédie expérimentale qu’il dénomme Citizendium, acronyme de «The Citizens' Compendium». Citizendium est une encyclopédie coopérative gratuite, tout comme Wikipédia, mais, d’après Larry Sanger, sans ses travers, à savoir le vandalisme, le manque de rigueur et l'utilisation d'un pseudonyme pour y contribuer. Les auteurs signent leurs articles de leur vrai nom et ces articles sont relus et corrigés par des «editors», âgés d'au moins 25 ans et titulaires d'une licence universitaire. De plus, des «constables» sont chargés de la bonne marche du projet et du respect du règlement. Le jour de son lancement, Citizendium compte 1.100 articles, 820 auteurs et 180 experts. L’encyclopédie comprend 11.800 articles en juillet 2009 et 15.000 articles en septembre 2010.

Wikipédia inspire bien d’autres encyclopédies participatives au fil des ans, par exemple l’Encyclopedia of Life, un projet global qui recense toutes les espèces animales et végétales connues. Wikipédia instaurera aussi de nouvelles règles pour tenter de résoudre les problèmes de vandalisme et de manque de rigueur inévitables pour un projet de cette échelle.

L’Encyclopedia of Life

L’Encyclopedia of Life débute en mai 2007 pour regrouper les connaissances existantes sur les espèces animales et végétales. Son but est d’être un «macroscope» permettant de déceler les grandes tendances à partir d’un stock d’informations considérable, à la différence du microscope permettant l’étude de détail. Les espèces connues seraient au nombre de 1,8 million en 2007, y compris les espèces en voie d’extinction, avec l’ajout de nouvelles espèces au fur et à mesure de leur identification.

Ce projet collaboratif est mené par plusieurs grandes institutions (Field Museum of Natural History, Harvard University, Marine Biological Laboratory, Missouri Botanical Garden, Smithsonian Institution, Biodiversity Heritage Library). Son financement initial est assuré par

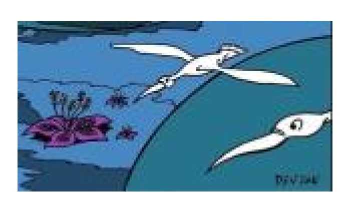
la MacArthur Foundation (10 millions de dollars US) et la Sloan Foundation (2,5 millions de dollars US).

Le directeur honoraire du projet est Edward Wilson, professeur émérite à Harvard, qui fut le premier à émettre le vœu d'une telle encyclopédie dans un essai daté de 2002. Cinq ans plus tard, c'est chose possible grâce aux avancées technologiques récentes: outils logiciels permettant l'agrégation de contenu, mash-up (le fait de rassembler un contenu donné à partir de nombreuses sources différentes) et gestion de contenus à grande échelle.

Cette encyclopédie compte offrir une page web par espèce, et donc un portail commun à des millions de documents épars (textes, photos, cartes, bandes sonores, vidéos) disponibles en ligne et hors ligne.

Outil d'apprentissage et d'enseignement pour une meilleure connaissance de notre planète, l'encyclopédie est à destination de tous (scientifiques, enseignants, étudiants, scolaires, médias, décideurs, grand public). Les contributions extérieures sont bienvenues, le contenu étant ensuite validé (ou non) par des experts. La version anglaise est traduite en plusieurs langues par des organismes partenaires.

@@@@@@@



13. La licence Creative Commons

Créée en 2001 à l'initiative de Lawrence “Larry” Lessig, la licence Creative Commons a pour but de favoriser la diffusion et la réutilisation d'œuvres numériques tout en protégeant le droit d'auteur. De nombreux auteurs et autres créateurs souhaitent en effet utiliser le réseau de diffusion qu’offre l’internet tout en contrôlant les droits sur leurs œuvres, mais le copyright traditionnel leur semble trop restrictif. Creative Commons fête ses 20 ans en juin 2021 avec 2 milliards d’œuvres sous licence Creative Commons.

***

Le copyleft, la GPL et la GFDL

Bien avant la licence Creative Commons, le copyleft est créé dès 1984 par Richard Stallman, ingénieur informatique au MIT (Massachusetts Institute of Technology) et ardent défenseur du logiciel libre au sein de la Free Software Foundation (FSF). Conçu à l’origine pour les logiciels, le copyleft est formalisé par la GPL (General Public License) et étendu ensuite à toute oeuvre de création.

Le copyleft inclut un «copyright» affirmant le droit d'auteur, mais donne à l’usager le droit de librement redistribuer le document et de le modifier. L’usager s’engage toutefois à ne revendiquer ni le travail original, ni les œuvres dérivées, qui sont elles aussi placées d’emblée sous licence GPL.

La licence GPL se décline en plusieurs variantes. Créée en 2000, la GFDL (GNU Free Documentation License) est la licence utilisée pour la documentation accompagnant les logiciels libres (manuels, livres et autres documents). La GFDL est ensuite utilisée pour toute œuvre documentaire, y compris les dictionnaires et les encyclopédies en ligne. Wikipédia par exemple utilise la GFDL avant d’utiliser la licence Creative Commons.

La licence Creative Commons

La licence Creative Commons est conçue en 2001 à l’initiative de Lawrence «Larry» Lessig. Ce juriste, qui est professeur de droit à la Stanford Law School (Californie), défend ardemment un internet créatif sur lequel les œuvres pourraient être non seulement diffusées mais aussi réutilisées lorsque les auteurs donnent leur accord.

L'organisme du même nom propose six licences-type, qui sont des contrats flexibles de droit d'auteur compatibles avec une diffusion sur l'internet. Ces autorisations non exclusives permettent aux titulaires des droits d'autoriser (ou non) la diffusion et l’utilisation de leurs œuvres (texte, photo, vidéo, film, musique, site web, etc.) pour une exploitation commerciale

ou pour des œuvres dérivées, par exemple. Rédigées par des juristes, ces licences-type sont accompagnées d'un résumé court accessible au commun des mortels.

Comment choisir une licence?

Avant de publier son œuvre sur l'internet, l'auteur choisit la licence adéquate en fonction de ses souhaits -- utilisation commerciale ou non, possibilité d'œuvre dérivée ou non, utilisation de la même licence ou non pour les œuvres dérivées, etc. -- et appose sur son œuvre un lien vers la licence correspondante sur le site de Creative Commons, ainsi que le logo correspondant (petit ou grand modèle).

Six licences sont disponibles:

1. La licence CC BY (Attribution) est la plus large puisqu'elle requiert seulement la mention de l'auteur, du titre et de la source en cas de réutilisation de l'œuvre.

2. La licence CC BY-SA (Attribution-ShareAlike) requiert la même licence en cas de réutilisation de l'œuvre.

3. La licence CC BY-ND (Attribution-NoDerivs) interdit les œuvres dérivées en cas de réutilisation de l'œuvre.

4. La licence CC BY-NC (Attribution-NonCommercial) interdit un usage commercial en cas de réutilisation de l'œuvre.

5. La licence CC BY-NC-SA (Attribution-NonCommercial-ShareAlike) interdit un usage commercial et requiert la même licence en cas de réutilisation de l'œuvre.

6. La licence CC BY-NC-ND (Attribution-NonCommercial-NoDerivs) interdit un usage commercial et des œuvres dérivées en cas de réutilisation de l'œuvre.

De la version 1.0 à la version 4.0

Après une version 1.0 (publiée en décembre 2002) suivie d'une version 2.0 (publiée en mai 2004), la version 3.0 (publiée en février 2007) instaure une licence internationale et la compatibilité avec d'autres licences similaires, dont le copyleft et la GPL.

Suite à un sommet global organisé en septembre 2011 auquel s'ajoutent deux années de discussions publiques documentées sur le web, Creative Commons publie la version 4.0 de ses six licences en novembre 2013.

La version 4.0 (disponible en 23 langues) supprime les versions nationales pour favoriser la licence internationale. Elle requiert aussi que les modalités d'utilisation des œuvres soient mieux indiquées sur les œuvres elles-mêmes, avec un lien systématique vers le descriptif de la licence sur le site de Creative Commons, une étape souvent oubliée par les usagers.

Qui utilise une licence Creative Commons?

Le premier éditeur à utiliser une licence Creative Commons est O'Reilly Media. Fondé par Tim O'Reilly en 1978, O'Reilly Media est un éditeur réputé de manuels informatiques et de livres sur les technologies.

Dans un premier temps, alors que la Creative Commons n'existe pas encore, l'éditeur offre une formule de «copyright ouvert» pour les auteurs qui le souhaitent ou pour des projets collectifs, et il permet d'emblée à ses auteurs de diffuser simultanément une version numérique gratuite de leurs livres.

En 2003, il privilégie le Creative Commons Founders' Copyright, à savoir un copyright ramené à 14 ou 28 ans après la mort de l'auteur (et qui était la norme il y a bien longtemps), et il ajoute une licence Creative Commons à tous ses livres.

Une licence Creative Commons est utilisée aussi par Wikipédia. Après avoir utilisé une licence GFDL, Wikipédia utilise une licence CC BY-SA pour tous ses articles, rédigés par des milliers de contributeurs dans des dizaines de langues différentes, auxquels s'ajoutent tous ceux qui fournissent des illustrations (images, photos, dessins, cartes, graphiques, etc.). Les textes et les illustrations peuvent être réutilisés par tous dans la mesure où la même licence est utilisée.

La Public Library of Science (PLOS) utilise également une licence Creative Commons pour les articles de ses sept revues scientifiques et médicales gratuites de haut niveau créées entre 2003 et 2007 (PLOS Biology, PLOS Medicine, PLOS Genetics, PLOS Computational Biology, PLOS Pathogens, PLOS Neglected Tropical Diseases et PLOS ONE).

PLOS utilise une licence CC BY, la plus large qui soit. Tous les articles peuvent être librement diffusés et réutilisés ailleurs, y compris pour des traductions, la seule contrainte étant la mention des auteurs, du titre et de la source. Ces revues gratuites ne tardent pas à rivaliser avec les meilleures revues scientifiques payantes (et hors de prix), avec l'avantage énorme d'être accessibles aux pauvres comme aux riches où qu'ils soient sur notre planète.

Suivent aussi des milliers d'auteurs, traducteurs, illustrateurs, musiciens, enseignants, chercheurs et autres créateurs partageant leurs travaux sur la toile, et pour certains dans un esprit militant pour contribuer à un monde meilleur.

Une licence Creative Commons est utilisée pour un million d'œuvres en 2003, 4,7 millions d'œuvres en 2004, 20 millions d'œuvres en 2005, 50 millions d'œuvres en 2006, 90 millions d'œuvres en 2007, 130 millions d'œuvres en 2008, 400 millions d'œuvres en 2010, 882 millions d'œuvres en 2014, 1,6 milliard d'œuvres en 2020 et 2 milliards d'œuvres en juin 2021, un beau chiffre pour célébrer ses 20 ans.

@@@@@@@

 14. De Google Print à Google Books

Google lance Google Print en mai 2005 en partenariat avec des éditeurs et des bibliothèques, puis le suspend trois mois plus tard suite à un conflit avec les associations américaines d'auteurs et d'éditeurs. Le programme reprend en août 2006 sous le nom de Google Books (Google Livres) pour numériser les milliers de livres de bibliothèques partenaires, y compris les livres sous droits sans l'autorisation préalable des ayants droit. Google invoque en effet le «fair use» (droit de citation). Après une saga judiciaire de près de dix ans, la justice américaine tranche en octobre 2015 en faveur de Google.

***

Google Print

Google décide de mettre son expertise au service du livre et lance Google Print en mai 2005. Ce lancement officiel est précédé de deux étapes.

En octobre 2004, Google lance la première partie de son programme Google Print en partenariat avec les éditeurs pour pouvoir consulter à l'écran des extraits de livres, puis commander ces livres auprès d'une librairie en ligne.

En décembre 2004, Google lance la deuxième partie de son programme Google Print, cette fois à destination des bibliothèques. Google numérise les livres appartenant à plusieurs grandes bibliothèques partenaires, à commencer par la bibliothèque de l'Université du Michigan (dans sa totalité, à savoir 7 millions d'ouvrages), les bibliothèques des universités de Harvard, de Stanford et d'Oxford, et la bibliothèque publique de New York (New York Public Library).

Le coût estimé au départ se situe entre 150 et 200 millions de dollars US, avec la numérisation de 10 millions de livres sur six ans et un chantier d'une durée totale de dix ans.

En août 2005, soit trois mois après son lancement, Google Print est suspendu pour une durée indéterminée suite à un conflit grandissant avec les associations américaines d'auteurs et d'éditeurs, celles-ci reprochant à Google de numériser des livres sans l'accord préalable des ayants droit.

Google Books (Google Livres)

Le programme reprend en août 2006 sous le nom de Google Books (Google Livres), qui permet de rechercher les livres par date, titre ou éditeur. La numérisation des fonds de grandes bibliothèques se poursuit, tout comme des partenariats avec les éditeurs qui le souhaitent.

Les livres libres de droit sont consultables à l'écran en texte intégral. Leur contenu est copiable et l'impression est possible page à page. Ces livres sont téléchargeables sous forme de fichiers PDF imprimables dans leur entier. Les liens publicitaires présents sur toute page de Google sont situés en haut et à droite de l'écran.

Le conflit avec les associations d'auteurs et d'éditeurs se poursuit lui aussi, puisque Google continue de numériser des livres sous droits sans l'autorisation préalable des ayants droit, en invoquant le «fair use» (droit de citation) pour présenter des extraits sur le web. L'Authors Guild et l'Association of American Publishers (AAP) invoquent pour leur part le non respect de la législation américaine du copyright pour attaquer Google en justice.

Fin 2006, d'après le buzz médiatique, Google scannerait 3.000 livres par jour, ce qui représenterait un million de livres par an. Le coût estimé serait de 30 dollars par livre. Google Books comprendrait déjà 3 millions de livres. Tous chiffres à prendre avec précaution, Google ne communiquant pas de statistiques à ce sujet.

Les bibliothèques partenaires

À l'exception de la bibliothèque publique de New York (New York Public Library), les premières bibliothèques numérisées sont toutes des bibliothèques universitaires (Harvard, Stanford, Michigan, Oxford, Californie, Virginie, Wisconsin-Madison, Complutense de Madrid), auxquelles s'ajoutent début 2007 les bibliothèques des universités de Princeton et du Texas (Austin), la Biblioteca de Catalunya (Catalogne, Espagne) et la Bayerische Staatbibliothek (Bavière, Allemagne).

En mai 2007, Google annonce la participation de la première bibliothèque francophone, la Bibliothèque cantonale et universitaire (BCU) de Lausanne (Suisse), avec la numérisation de 100.000 titres en français, en allemand et en italien publiés entre le 17e et le 19e siècle. Suit un partenariat avec la Bibliothèque municipale de Lyon (France) signé en juillet 2008 pour numériser 500.000 livres.

En novembre 2008, Google Books comprend 7 millions d'ouvrages numérisés, en partenariat avec 24 bibliothèques et 2.000 éditeurs. Les 24 bibliothèques partenaires sont 16 bibliothèques aux États-Unis, deux bibliothèques en Espagne, et une bibliothèque dans chacun de ces pays: Allemagne, Belgique, France, Japon, Royaume-Uni et Suisse.

Le conflit avec les ayants droit

En octobre 2008, après trois ans de conflit, Google tente de mettre fin aux procès intentés par les associations d'auteurs et d'éditeurs en proposant un accord basé sur un partage des revenus générés par Google Books ainsi qu'un large accès aux ouvrages épuisés, tout comme le paiement de 125 millions de dollars US à l'Authors Guild et à l'Association of American Publishers (AAP) pour clôturer définitivement ce conflit.

Cet accord permettrait à Google de proposer de plus larges extraits de livres (jusqu'à 20% d'un même ouvrage) avec un lien commercial pour acheter une copie (numérique ou non) de l'œuvre. Les ayants droit auraient la possibilité de demander le retrait de leurs livres. Les bibliothèques universitaires et publiques des États-Unis pourraient accéder à un portail gratuit donnant accès aux millions de livres épuisés. Un abonnement permettrait aux universités et aux écoles de consulter les collections des bibliothèques les plus renommées.

Mais cette proposition d'accord est très critiquée puis refusée. La justice américaine tranche toutefois en faveur de Google en octobre 2015 après une saga judiciaire de près de dix ans. Google a toujours invoqué le «fair use» (droit de citation) pour les livres sous droits et c'est l'argument qui prévaut pour mettre fin au conflit.

@@@@@@@



15. L'Internet Archive, une bibliothèque planétaire

L’Internet Archive est fondée en avril 1996 par Brewster Kahle à San Francisco (Californie) pour préserver un historique de l’internet. D’abord «bibliothèque de l’internet» avec la Wayback Machine, qui permet de voir l’évolution d’un site web depuis 1996, elle devient ensuite «une bibliothèque numérique à but non lucratif destinée à procurer un accès universel au savoir humain». Elle propose deux millions de livres numérisés en mars 2010. Dix ans plus tard, elle est la plus grande bibliothèque numérique publique mondiale.

***

L’archivage de l’internet

L’Internet Archive a d’abord pour but de constituer, stocker, préserver et gérer une archive de l’internet, en stockant la totalité du web tous les deux mois pour offrir un outil de travail aux universitaires, chercheurs et historiens, et préserver un historique de l’internet pour les générations présentes et futures.

Comme expliqué à l’époque sur le site, de tout temps les sociétés ont voulu préserver leur culture et leur héritage pour les générations présentes et futures. Les bibliothèques ont eu pour vocation de conserver les traces écrites de ces cultures et de ces héritages, et d’en procurer l’accès au grand public et aux chercheurs. Il paraît donc essentiel qu’elles étendent leur mission aux nouvelles technologies.

Paradoxalement, le travail de sauvegarde a souvent été mal fait au début du 20e siècle. Nombreux ont été les premiers films qui ont été recyclés -- et donc définitivement perdus -- pour récupérer la couche d’argent présente sur la pellicule. Nombre d’émissions de radio et de télévision n’ont pas été conservées. Il importe donc de ne pas reproduire la même erreur pour l’internet, et particulièrement pour le web, un nouveau médium dont l’immense portée est encore méconnue en 1996, cinq ans après son invention. C’est la raison d’être de l’Internet Archive.

La Wayback Machine

En octobre 2001, avec 30 milliards de pages archivées, l'Internet Archive met ses archives en accès libre grâce à la Wayback Machine, qui permet à tout un chacun de voir l'historique d'un site web, à savoir la présentation et le contenu d'un site web donné tous les deux mois depuis avril 1996 (date de la création de l'Internet Archive).

En 2004, les archives du web représentent plus de 300 To (téra-octets) de données, avec une croissance de 12 To par mois. Le nombre de pages web visibles avec la Wayback Machine est de 65 milliards (provenant de 50 millions de sites web) en décembre 2006, 85 milliards en mai 2007 et 150 milliards en mars 2010.

Un projet de bibliothèque numérique

En 2000, l'Internet Archive débute la constitution de collections numériques, en hébergeant notamment les livres du Million Book Project (10.520 livres en avril 2005), tout comme des archives de films de la période 1903-1973, des archives de concerts live récents, des archives de logiciels, des archives d'images et de vidéos, les sites relatifs au 11 septembre 2001, les sites relatifs aux élections américaines de 2000 (présidentielles) et 2002 (Chambre des représentants), les sites relatifs aux pionniers du web, etc. Toutes ces collections sont en consultation libre.

L'Internet Archive pense aussi qu'une bibliothèque numérique à vocation mondiale ne doit pas être liée à des enjeux commerciaux. Elle veut éviter les travers de Google Books, à savoir la numérisation des livres sous droits sans l'accord préalable des ayants droit, la recherche possible uniquement sur les outils de Google (Google Search et Google Chrome), l'existence de liens pour l'achat de ces livres et la présence de liens publicitaires en haut et à droite de l'écran.

L'Internet Archive lance donc l'Open Content Alliance (OCA) en octobre 2005 pour fédérer un grand nombre de partenaires (bibliothèques, universités, organismes gouvernementaux, associations, organismes culturels, sociétés informatiques). Cette bibliothèque de livres numérisés et de documents multimédia serait respectueuse du copyright des ayants droit et elle aurait un modèle ouvert, avec des collections consultables sur tout moteur de recherche.

La bibliothèque numérique

Les premiers organismes participants sont les bibliothèques des universités de Californie et de Toronto, les Archives européennes, les Archives nationales du Royaume-Uni, O'Reilly Media et les Prelinger Archives.

Seuls les livres appartenant au domaine public sont numérisés, pour éviter les problèmes de copyright auxquels se heurte Google. Les livres sous droits ne sont numérisés que si les ayants droit ont expressément donné leur accord. Les collections comprennent 100.000 livres numérisés en décembre 2006, avec 12.000 nouveaux livres par mois, et 200.000 livres numérisés en mai 2007.

En décembre 2006, l'Internet Archive reçoit une subvention d'un million de dollars US de la part de la Sloan Foundation pour numériser plusieurs collections: les collections complètes du Metropolitan Museum of Art (la totalité des livres et plusieurs milliers d'images), une collection de la Boston Public Library (les 3.800 livres de la bibliothèque personnelle de John

Adams, deuxième président des États-Unis), une collection de livres d'art du Getty Research Institute, une collection de la John Hopkins University (un ensemble de documents sur le mouvement anti- esclavagiste) et enfin une collection de l'Université de Californie à Berkeley (un ensemble de documents sur la ruée vers l'or).

Les 750.000 livres numérisés par Microsoft pour son projet Live Search Books (débuté en décembre 2006) sont versés dans les collections de l'Internet Archive en mai 2008 lorsque Microsoft met un terme à son projet.

L'Internet Archive offre un million de livres numérisés en décembre 2008 et deux millions de livres numérisés en mars 2010. Dix ans plus tard, elle est la plus grande bibliothèque numérique publique mondiale.

@@@@@@@



16. L’ebook vu par quelques pionniers

Voici le sentiment de plusieurs pionniers du livre numérique interviewés par courriel au fil des ans pour mes articles et livres sur le sujet: Michael Hart (Projet Gutenberg), John Mark Ockerbloom (The Online Book Page), Robert Beard (A Web of Online Dictionaries), Jean-Paul (Cotres.net), Nicolas Pewny (Le Choucas), Marc Autret (Indiscripts), Pierre Schweitzer (@folio), Denis Zwirn (Numilog) et Henri “Henk” Slettenhaar (Silicon Valley Association).

***

Michael Hart (Projet Gutenberg)

«Nous considérons le texte électronique comme un nouveau médium, sans véritable relation avec le papier. Le seul point commun est que nous diffusons les mêmes œuvres, mais je ne vois pas comment le papier peut concurrencer le texte électronique une fois que les gens y sont habitués, particulièrement dans les établissements d'enseignement.» (Michael Hart, fondateur du Projet Gutenberg et père de l’ebook, interviewé en août 1998)

Mark Ockerbloom (The Online Books Page)

«Je me suis passionné pour l'énorme potentiel qu'a l'internet de rendre la littérature accessible au plus grand nombre. (…) Je suis très intéressé par le développement de l'internet en tant que médium de communication de masse ces prochaines années. J'aimerais aussi rester impliqué dans la mise à disposition gratuite de livres sur l'internet, que ceci fasse partie intégrante de mon activité professionnelle, ou que ceci soit une activité bénévole menée sur mon temps libre.» (Mark Ockerbloom, bibliothécaire numérique et créateur de The Online Books Page, interviewé en septembre 1998)

Robert Beard (A Web of Online Dictionaries)

«En tant que professeur de langue, je pense que le web présente une pléthore de nouvelles ressources disponibles dans la langue étudiée, et disponibles 24 heures sur 24 et 7 jours sur 7. Aussi bien pour mes collègues que pour moi, et bien sûr pour notre établissement, l'internet nous permet aussi de publier pratiquement sans limitation. (...) L’internet nous offrira tout le matériel pédagogique dont nous pouvons rêver, y compris des notes de lecture, exercices, tests, évaluations et exercices interactifs plus efficaces dont certains sont

encore du domaine de l’avenir. Le web sera une encyclopédie du monde faite par le monde pour le monde.» (Robert Beard, professeur de langue et créateur de A Web of Online Dictionaries et yourDictionary.com, interviewé en octobre 1998)

Jean-Paul (Cotres.net)

«La navigation par hyperliens se fait en rayon (j'ai un centre d'intérêt et je clique méthodiquement sur tous les liens qui s'y rapportent) ou en louvoiements (de clic en clic, à mesure qu'ils apparaissent, au risque de perdre de vue mon sujet). Bien sûr, les deux sont possibles avec l'imprimé. Mais la différence saute aux yeux: feuilleter n'est pas cliquer. L'internet a donc changé mon rapport à l'écriture. (...) C'est finalement dans la publication en ligne (l'entoilage?) que j'ai trouvé la mobilité, la fluidité que je cherchais.» (Jean-Paul, auteur hypermédia et créateur du site Cotres.net, interviewé en juin 2000)

Nicolas Pewny (Le Choucas)

«Je vois le livre numérique du futur comme un "ouvrage total" réunissant textes, sons, images, vidéo, interactivité: une nouvelle manière de concevoir et d'écrire et de lire, peut-être sur un livre unique, sans cesse renouvelable, qui contiendrait tout ce que l'on a lu, unique et multiple compagnon. Utopique? Invraisemblable? Peut-être pas tant que cela!» (Nicolas Pewny, fondateur des éditions du Choucas puis consultant en édition électronique, interviewé en février 2003)

Marc Autret (Indiscripts)

«Il y a au moins deux axes qui émergent pour l’ebook: (a) une interface de lecture/consultation de plus en plus attractive et fonctionnelle (navigation, recherche, restructuration à la volée, annotations de l’utilisateur, quiz interactif...); (b) une intégration multimédia (vidéo, son, infographie animée, base de données...) désormais fortement couplée au web. Aucun livre physique n’offre de telles fonctionnalités. J’imagine donc l'ebook de demain comme une sorte de wiki cristallisé, empaqueté dans un format. Quelle sera alors sa valeur propre? Celle d'un livre: l'unité et la qualité du travail éditorial!» (Marc Autret, développeur, infographiste et créateur du site Indiscripts, interviewé en décembre 2006)

Pierre Schweitzer (@folio)

«La chance qu’on a tous est de vivre là, ici et maintenant cette transformation fantastique. Quand je suis né en 1963, les ordinateurs avaient comme mémoire quelques pages de caractères à peine. Aujourd’hui, mon baladeur de musique pourrait contenir des milliards de pages, une vraie bibliothèque de quartier. Demain, par l’effet conjugué de la loi de Moore et de l’omniprésence des réseaux, l’accès instantané aux œuvres et aux savoirs sera de mise. Le support de stockage lui-même n’aura plus beaucoup d’intérêt. Seules importeront les commodités fonctionnelles d’usage et la poétique de ces objets.» (Pierre Schweitzer,

concepteur du lecteur de textes @folio et du logiciel de remise en page Mot@Mot, interviewé en décembre 2006)

Denis Zwirn (Numilog)

«Le livre numérique n'est plus une question de colloque, de définition conceptuelle ou de divination par certains "experts": c'est un produit commercial et un outil au service de la lecture. (...) Il suffit de proposer des textes lisibles facilement sur les supports de lecture électronique variés qu'utilisent les gens, l'encre électronique pouvant progressivement envahir tous ces supports. Et de les proposer de manière industrielle. Ce n'est pas et ne sera jamais un produit de niche (les dictionnaires, les guides de voyage, les livres pour les non voyants...): c'est en train de devenir un produit de masse, riche de formes multiples comme l'est le livre traditionnel.» (Denis Zwirn, fondateur de la librairie numérique Numilog, interviewé en août 2007)

Henri “Henk” Slettenhaar (Silicon Valley Association)

«Je n'ai jamais aimé lire un livre sur un ordinateur ou sur un PDA. Maintenant, avec l’arrivée de tablettes comme le Kindle et l’iPad, je suis finalement devenu un lecteur de livres numériques. Je vois une expansion énorme avec l'arrivée de tablettes faciles à utiliser et avec un choix considérable de livres grâce au commerce électronique et à des sociétés comme Amazon. (…) J'utilise également des livres en ligne pour apprendre l'art de l'innovation!» (Henri “Henk” Slettenhaar, professeur des technologies de la communication et directeur de la Silicon Valley Association suisse, interviewé en juin 2011)

@@@@@@@



17. Un hommage aux bibliothécaires

Le/la bibliothécaire a toujours eu un rôle charnière dans le monde du livre, y compris numérique. Les bibliothèques sont largement couvertes par les médias, mais qu'en est-il de ceux et celles qui les font fonctionner? Une première version de ce texte fut publiée dans ActuaLitté en 2016. Voici une version revue et actualisée pour inclure les ultimes développements de la profession. L'auteure fut elle-même bibliothécaire avant d'être traductrice.

***

Le/la bibliothécaire officie souvent dans une bibliothèque renommée médiathèque, qui comprend non seulement une salle de référence et un service de prêt, mais aussi une vidéothèque, une artothèque, une salle d'exposition, une salle de conférences (avec heures du conte, rencontres, débats, signatures, concerts, et séances de yoga ou de tai-chi), un cyberespace et une bibliothèque numérique.

Et, dans le meilleur des cas, un véritable café plutôt qu'une simple machine à boissons sous l'escalier.

Ultime virage technologique, le/la bibliothécaire lit de plus en plus de livres numériques (appelés aussi ebooks) sur son ordinateur (portable ou non), son smartphone, sa liseuse ou sa tablette multifonctions. Avec plus ou moins de conviction les premiers temps, histoire d'être dans le coup, avant d'être totalement séduit par ces nouveaux supports.

Le/la bibliothécaire s'inquiète toutefois de voir l'ebook envahir le monde malgré son très jeune âge et menacer le livre imprimé et son modèle économique bien rodé. Après plus de cinq siècles de bons et loyaux services, le livre imprimé serait moribond et le support papier serait en phase terminale, selon quelques prophètes numériques.

Le/la bibliothécaire rassemble aussitôt tous les professionnels du livre (auteurs, traducteurs, illustrateurs, photographes, designers, éditeurs, libraires, diffuseurs, distributeurs, etc.) et des milliers de lecteurs et lectrices pour catalyser leur énergie vers un seul et unique noble but.

L'important est l'œuvre, explique-t-il/elle sans relâche, et non le support. Deux supports (imprimé et numérique) sont possibles. Aimer, c'est regarder dans la même direction.

Suite à quoi les smartphones, les liseuses et les tablettes multifonctions font bon ménage avec les livres imprimés et vont jusqu'à partager les mêmes stands dans les salons du livre -- en attendant d'occuper toute la place.

Contrairement à ce que suggèrent certaines mauvaises langues, le/la bibliothécaire n'a absolument pas le temps de lire au travail.

En plus de ses tâches habituelles (achat, catalogage, indexation, prêt, rangement, inventaire et bien sûr désherbage), il/elle alimente de manière continue son blog professionnel, son mur Facebook et son fil Twitter tout en gérant son compte Instagram et ses communications sur WhatsApp.

Lorsque les commentaires fusant en jet continu sous le dernier post de son blog ou de sa page Facebook se font un peu moins nombreux, il lui arrive même d'actualiser une notice de Wikipédia.

Certains deviendraient fous à faire autant de choses à la fois. Mais, à force de volonté personnelle doublée de séances de relaxation, le/la bibliothécaire a appris depuis longtemps à gérer son stress, tout en luttant au sein de son syndicat pour sauver son emploi et celui de ses collègues, et demander une augmentation de salaire dans la foulée.

La nuit, en cas de sommeil agité, le/la bibliothécaire non technophile rêve de l'époque où l'internet n'avait pas encore envahi la planète en général, et les bibliothèques en particulier.

Pendant ce temps, le/la bibliothécaire technophile rêve au (véritable) papier électronique, toujours dans les éprouvettes des chercheurs malgré un lancement prochain régulièrement reporté d'année en année.

Après cette courte digression nocturne, on terminera en soulignant avec force le rôle pivot du bibliothécaire -- en version classique et en version numérique -- dans un monde du livre en pleine mutation.

Le/la bibliothécaire était déjà là pour mettre de l'ordre dans ce que Gutenberg et ses successeurs imprimaient. Le/la bibliothécaire était encore là pour organiser les premières bibliothèques municipales. Le/la bibliothécaire était toujours là pour promouvoir la lecture publique auprès de générations de petits et grands lecteurs.

Le/la bibliothécaire gère désormais non seulement des documents imprimés, mais aussi des CD, des DVD, des pages web, des fichiers texte, des fichiers audio, des fichiers vidéo et des fichiers multimédia, et bien sûr des livres numériques en quantité dans de nombreux formats.

Vu le nombre de documents, la multiplicité des supports et l'omniprésence des réseaux sociaux, sa tâche devient de plus en plus rude, et la pandémie en cours n'arrange rien.

Mais, depuis cinq siècles que le/la bibliothécaire est d'attaque sur tous les fronts, un pari de plus (et un salaire en conséquence) ne lui fait pas peur.

@@@@@@@



18. Une chronologie de 1971 à nos jours

Cette chronologie couvre le livre numérique sous tous ses aspects: livres gratuits, livres payants, bibliothèques numériques, librairies en ligne, éditeurs en ligne, formats numériques, logiciels de lecture, PDA, smartphones, liseuses, tablettes, dictionnaires, encyclopédies, projets innovants, etc.

***
[année-mois]

1971-07 > Michael Hart saisit l'eText #1 (qui deviendra l'eBook #1) le 4 juillet 1971 sur le clavier de son terminal et le fait savoir aux 100 usagers du réseau pré-internet. Six usagers le téléchargent. Le Projet Gutenberg est né.

1974 > Vinton Cerf et Robert Kahn créent les protocoles de communication de l'internet.

1983 > Après avoir été un réseau gouvernemental et académique américain, l'internet prend son envol à l'international.

1984 > Richard Stallman fonde la Free Software Foundation (FSF), qui institue le copyleft et la licence GPL (General Public License) pour les logiciels.

1991 > Tim Berners-Lee invente le World Wide Web et donne son invention au monde.

1993 > John Mark Ockerbloom crée The Online Books Page pour offrir un point commun aux livres numériques en accès libre sur le web.

1993-04 > L'Association des bibliophiles universels (ABU) crée la première bibliothèque numérique française.

1993-06 > Adobe lance le format PDF (Portable Document Format), tout comme l'Acrobat Reader (pour lire les PDF) et l'Adobe Acrobat (pour créer les PDF).

1994-02 > La bibliothèque publique d'Helsinki (Finlande) est la première bibliothèque à créer un site web.

1995 > L'Institut national de la langue française (INaLF) crée sa base textuelle Frantext en accès payant.

1997-04 > PARC (Palo Alto Research Center), le centre Xerox de la Silicon Valley, conçoit le Gyricon, une technologie d'encre électronique.

1995-07 > Amazon.com ouvre ses portes virtuelles en tant que première grande librairie en ligne américaine.

1995-09 > Des bibliothèques nationales européennes créent le site web Gabriel en tant que point d'accès commun à leurs services.

1996-03 > Le premier PDA (Personal Digital Assistant) du marché est le Palm Pilot.

1996-04 > Brewster Kahle fonde l'Internet Archive pour archiver l'internet à l'intention des générations présentes et futures.

1996-06 > Isabelle Aveline crée Zazieweb, un site indépendant couvrant l'actualité du livre.

1996-07 > Olivier Gainon fonde CyLibris, le premier éditeur électronique commercial de langue française.

1996-10 > Pierre Schweitzer conçoit le baladeur de textes @folio.

1996-11 > Nicolas Pewny crée le site web du Choucas, sa maison d'édition.

1997 > Hachette met en ligne son Dictionnaire universel francophone, avec accès libre et gratuit.

1997-01 > Gabriel devient le portail web des bibliothèques nationales européennes.

1997-03 > La chaîne de librairies Barnes & Noble ouvre une première librairie en ligne.

1997-04 > Des chercheurs du Media Lab du MIT créent la société E Ink pour développer une technologie d'encre électronique.

1997-04 > Le web mondial compte un million de sites.

1997-10 > La Bibliothèque nationale de France (BnF) crée sa bibliothèque numérique Gallica avec 3.000 ouvrages du 19e siècle.

1997-12 > L'internet mondial compte 70 millions d'usagers, soit 1,7% de la population mondiale.

1998 > NuvoMedia lance le Rocket eBook, la première liseuse du marché, avec une capacité de dix livres.

1998 > SoftBook Press lance le SoftBook Reader, la deuxième liseuse du marché.

1998-05 > Les éditions 00h00 sont les premières à vendre des livres numériques.

1998-06 > L'Open eBook Initiative est instituée pour créer un standard de livre numérique.

1998-10 > Jean-Paul crée le site hypermédia Cotres.net pour expérimenter de nouvelles formes d'écriture utilisant l'hyperlien.

1999 > Michael Kellogg crée WordReference.com pour proposer des dictionnaires de langues gratuits et des forums de discussion pour les linguistes.

1999-09 > Le format Open eBook (OeB) est créé en tant que standard du livre numérique (à côté du format pionnier PDF).

1999-12 > Denis Zwirn fonde la librairie Numilog, consacrée d'emblée à la vente de livres numériques.

1999-12 > L'Encyclopaedia Britannica met en ligne sa version numérique Britannica.com, d'abord gratuite puis payante.

1999-12 > L'Encyclopaedia Universalis propose sa version en ligne payante, avec de nombreux articles en accès libre.

1999-12 > WebEncyclo (éditions Atlas) est la première grande encyclopédie francophone en ligne gratuite.

2000 > La licence GFDL (GNU Free Documentation License) est une variante de la licence GPL (General Public License) pour les manuels de logiciels libres.

2000-01 > Gemstar rachète les sociétés Nuvomedia et SoftBook Press (auteures des premières liseuses du marché) pour lancer ses propres liseuses l'année suivante.

2000-01 > L'Open eBook Forum (OeBF) est fondé pour promouvoir le format OeB (Open eBook) en tant que standard du livre numérique.

2000-03 > L'Oxford University Press (OUP) crée une version en ligne de l'Oxford English Dictionary (OED), avec consultation payante.

2000-03 > Stephen King est le premier auteur de best-sellers à distribuer sa nouvelle «Riding the Bullet» uniquement sur l'internet.

2000-03 > Thierry Brethes et Nathalie Ting créent la société Mobipocket, spécialisée d'emblée dans les livres pour PDA, et lancent le logiciel de lecture Mobipocket Reader.

2000-03 > L'internet compte 300 millions d'usagers, soit 5% de la population mondiale.

2000-04 > Microsoft lance son propre PDA, le Pocket PC, tout comme le logiciel de lecture Microsoft Reader.

2000-07 > Stephen King auto-publie les premiers chapitres de son roman épistolaire inédit «The Plant» sur un site dédié, une expérience qui dure six mois.

2000-08 > Barnes & Noble.com crée son eBookStore pour vendre des livres numériques.

2000-08 > Le logiciel de lecture Microsoft Reader est disponible pour toute plateforme Windows.

2000-09 > Le Grand dictionnaire terminologique (GDT) du Québec est un dictionnaire en ligne gratuit bilingue français-anglais.

2000-10 > Charles Franks crée la plateforme Distributed Proofreaders pour la numérisation et la relecture des livres numériques du domaine public, y compris pour le Projet Gutenberg.

2000-10 > La société Franklin lance l'eBookMan, son propre PDA multimédia.

2000-10 > Gemstar lance ses propres liseuses sous le nom de Gemstar eBook, une expérience qui dure trois ans.

2000-11 > Le romancier espagnol Arturo Pérez-Reverte vend son nouveau roman «El Oro del Rey» en version numérique avant de le vendre en version imprimée.

2000-11 > Le romancier anglais Frederick Forsyth vend une nouvelle policière inédite chez l'éditeur électronique Online Originals, une expérience qui ne dure pas.

2000-11 > La société Nokia lance le Nokia 9210, premier smartphone du marché.

2000-12 > Amazon lance son eBookStore pour vendre des livres numériques.

2000-12 > Des chercheurs de PARC (Palo Alto Research Center) créent la société Gyricon Media pour développer une technologie d'encre électronique.

2001 > Lawrence "Larry" Lessig conçoit la licence Creative Commons et fonde l'organisme du même nom.

2001-01 > Jimmy Wales et Larry Sanger fondent Wikipédia, grande encyclopédie collaborative gratuite en ligne.

2001-01 > Le Cybook est la première liseuse européenne, créée par la société Cytale et reprise deux ans plus tard par la société Bookeen.

2001-03 > La société Palm lance le logiciel de lecture Palm Reader pour le Palm Pilot, son PDA.

2001-03 > Le logiciel de lecture Mobipocket Reader est intégré au Palm Pilot.

2001-04 > Un Seybold Report dénombre 17 millions de PDA dans le monde pour seulement 100.000 liseuses.

2001-05 > Le logiciel de lecture Acrobat Reader est disponible non seulement pour ordinateur mais aussi pour le Palm Pilot.

2001-09 > Yahoo! lance son eBookStore pour vendre des livres numériques.

2001-10 > Avec 30 milliards de pages archivées, l'Internet Archive lance la Wayback Machine pour voir un site web à différentes dates à compter de 1996.

2002 > 23 millions de Palm Pilot ont été vendus dans le monde depuis ses débuts en mars 1996 en tant que premier PDA du marché.

2002-03 > L'Oxford University Press (OUP) crée l'Oxford Reference Online (ORO), une vaste encyclopédie conçue directement pour le web, avec abonnement payant.

2002-07 > La société E Ink présente le prototype du futur écran E Ink pour liseuse, qui sera commercialisé deux ans plus tard.

2002-07 > Pierre Schweitzer crée la start-up iCodex pour développer son baladeur de textes @folio.

2002-12 > Les premières licences-type Creative Commons sont publiées par l'organisme du même nom.

2003-03 > Le romancier brésilien Paulo Coelho diffuse plusieurs de ses best-sellers gratuitement au format PDF.

2003-05 > L'Acrobat Reader (pour les PDF standard) et l'Adobe eBook Reader (pour les PDF sous droits) fusionnent pour donner l'Adobe Reader.

2003-12 > Adobe ouvre sa propre librairie numérique, le Digital Media Store.

2003-12 > Une licence Creative Commons est utilisée par un million d'œuvres sur l'internet.

2004 > La société E Ink commercialise les premiers écrans de liseuses utilisant la technologie du même nom.

2004-01 > La European Library (Bibliothèque européenne) succède à Gabriel en tant que portail des bibliothèques nationales européennes.

2004-04 > Sony lance au Japon sa première liseuse LIBRIe, qui est aussi la première liseuse du marché avec un écran E Ink et non un écran LCD (à cristaux liquides).

2004-10 > Google crée Google Print à l'intention des éditeurs et des bibliothèques avant de cesser le projet trois mois plus tard puis de le rebaptiser Google Books l'année suivante.

2005-04 > Amazon.com rachète la société Mobipocket – format, logiciel et livres – en prévision du lancement du Kindle deux ans plus tard.

2005-04 > L'IDPF (International Digital Publishing Forum) succède à l'OeBF (Open eBook Forum) pour gérer le format OeB (Open eBook).

2005-10 > L’Internet Archive fonde l’Open Content Alliance (OCA) pour fédérer de nombreux organismes dans le but de proposer une bibliothèque numérique mondiale.

2005-12 > L’internet mondial compte un milliard d’usagers, soit 15,7% de la population mondiale.

2006 > Le smartphone progresse, avec 90 millions de smartphones pour un milliard de téléphones portables sur notre planète.

2006-03 > La Commission européenne débute son projet de Bibliothèque numérique européenne, qui deviendra Europeana deux ans plus tard.

2006-08 > Google lance Google Books en remplacement de Google Print et entame une saga judiciaire de plusieurs années avec les associations d’auteurs et d’éditeurs.

2006-10 > Sony lance le Sony Reader, sa deuxième liseuse après le LIBRIe au Japon, là aussi avec un écran E Ink.

2006-11 > Le web mondial compte cent millions de sites.

2006-12 > Gallica, bibliothèque numérique de la Bibliothèque nationale de France, offre 90.000 ouvrages et 80.000 images libres de droits.

2007-02 > Creative Commons publie les versions 3.0 de ses six licences, avec licence internationale et compatibilité avec d’autres licences similaires, dont le copyleft et la GPL.

2007-05 > L’Encyclopedia of Life est créée en tant qu’encyclopédie collaborative en ligne visant à répertorier toutes les espèces animales et végétales connues.

2007-06 > Apple lance son propre smartphone, l’iPhone, qui devient un produit phare pour toute une génération.

2007-09 > L’International Digital Publishing Forum (IDPF) publie le format EPUB (Electronic PUBlication), qui remplace le format OeB (Open eBook).

2007-11 > Amazon.com lance sa propre liseuse, le Kindle.

2008-02 > Nicolas Gary crée la revue littéraire en ligne ActuaLitté, qui devient vite incontournable dans le monde du livre.

2008-07 > Le format PDF devient un standard ouvert et une norme ISO (ISO 32000-1:2008).

2008-11 > La Commission européenne crée Europeana en tant que grande bibliothèque numérique publique européenne.

2009-01 > Amazon rachète la société Audible.com et sa collection de livres, journaux et magazines audionumériques.

2009-01 > La librairie numérique Numilog propose un catalogue de 50.000 titres à la vente.

2009-02> Google Books propose un portail spécifique pour les livres numériques pour smartphone.

2009-05 > Amazon lance le Kindle DX avec un écran plus grand pour la lecture de journaux et magazines.

2009-05 > Marc Autret, développeur et infographiste, crée le site bilingue français-anglais Indiscripts, un laboratoire de scripts InDesign.

2009-11 > La chaîne de librairies américaine Barnes & Noble lance sa propre liseuse, le Nook.

2010-03 > Gallica offre un million de documents tous supports confondus.

2010-04 > Apple lance l’iPad, sa tablette multifonctions.

2010-05 > L’International Digital Publishing Forum (IDPF) publie l’EPUB2, à savoir la version 2 du format EPUB.

2010-10 > Distributed Proofreaders comptabilise 18.000 titres numérisés et relus pour le Projet Gutenberg.

2010-12 > Une licence Creative Commons est utilisée par 400 millions d’oeuvres sur l’internet.

2011-01 > Wikipédia fête ses 10 ans avec 17 millions d’articles dans 270 langues.

2011-03 > L’internet mondial compte 2 milliards d’usagers, soit 30,2% de la population mondiale.

2011-07 > Le Projet Gutenberg fête ses 40 ans avec une collection de 36.000 ebooks.

2011-09 > Michael Hart -- fondateur du Projet Gutenberg et père de l’ebook -- décède dans l’Illinois à l’âge de 64 ans après avoir consacré toute sa vie à son projet.

2011-10 > L’International Digital Publishing Forum (IDPF) publie l’EPUB3, à savoir la version 3 du format EPUB.

2013-11 > Creative Commons publie les versions 4.0 de ses six licences, qui sont d’emblée des licences internationales.

2014-12 > Une licence Creative Commons est utilisée par 882 millions d’œuvres sur l’internet.

2015-03 > L’internet mondial compte trois milliards d’usagers, soit 42,3% de la population mondiale.

2015-04 > The Online Books Page recense deux millions de livres en accès libre sur le web.

2015-05 > Le web mondial compte près d’un milliard de sites.

2015-07 > Distributed Proofreaders comptabilise 30.000 titres numérisés et relus pour le Projet Gutenberg.

2015-09 > Le Projet Gutenberg offre 50.000 livres numériques gratuits, pour la plupart du domaine public.

2017-01 > L’International Digital Publishing Forum (IDPF) est intégré au World Wide Web Consortium (W3C) qui poursuit le développement de l’EPUB.

2020-10 > Distributed Proofreaders fête ses 20 ans avec près de 40.000 titres numérisés et relus pour le Projet Gutenberg.

2021-01 > L’internet mondial compte 4,6 milliards d’usagers, soit 59,5% de la population mondiale.

2021-01 > Wikipédia fête ses 20 ans et sa croissance est toujours exponentielle.

2021-06 > Creative Commons fête ses 20 ans avec 2 milliards d’œuvres sous licence Creative Commons sur l’internet.

2021-07 > Le Projet Gutenberg fête ses 50 ans le 4 juillet 2021 avec plus de 65.000 livres numériques gratuits dans 60 langues et des dizaines de milliers de téléchargements par jour.

@@@@@@@